

AD-753 125

REPRESENTATION AND ESTIMATION OF CYCLO-
STATIONARY PROCESSES

William Allen Gardner

Massachusetts University

Prepared for:

Air Force Office of Scientific Research

August 1972

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AFOSR - TR - 72 - 2378

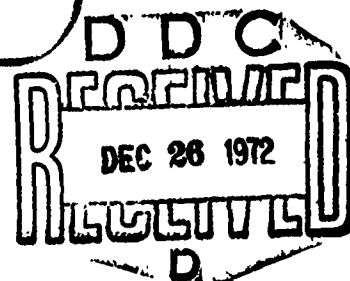
AD753125



Engineering Research Institute

UNIVERSITY OF MASSACHUSETTS

AMHERST, MASSACHUSETTS



Approved for public release;
distribution unlimited.

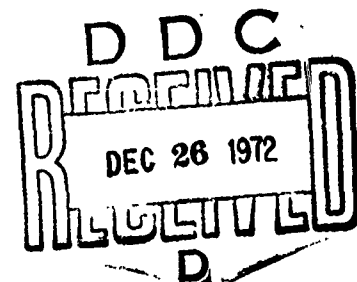
REPRESENTATION AND ESTIMATION OF
CYCLOSTATIONARY PROCESSES

by

William Allen Gardner

M.S., Stanford University

Ph.D., University of Massachusetts



ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 10/10/01 BY 60322

Amherst, Massachusetts

August 1972

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Engineering Research Institute University of Massachusetts Amherst, Massachusetts		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	
		2b. GROUP	
3. REPORT TITLE REPRESENTATION AND ESTIMATION OF CYCLOSTATIONARY PROCESSES			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific Interim			
5. AUTHOR(S) (First name, middle initial, last name) William A. Gardner			
6. REPORT DATE August 1972		7a. TOTAL NO. OF PAGES 353	7b. NO. OF FIGURES 63
8a. CONTRACT OR GRANT NO. AFOSR 71-2111		8b. ORIGINATOR'S REPORT NUMBER(S)	
b. PROJECT NO. 9769			
c. 61102F		8b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d. 681304		AFOSR - TR - 72 - 2378	
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES TECH., OTHER		12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research/NM 1400 Wilson Boulevard Arlington, Virginia 22209	
13. ABSTRACT Random signal processes which have been subjected to some form of repetitive operation such as sampling, scanning or multiplexing will usually exhibit statistical properties which vary periodically with time. In many cases, the repetitive operation is introduced intentionally to put the signal in a format which is easily manipulated and which preserves the time-position integrity of the events which the signal is representing. Familiar examples are radar antenna scanning patterns, raster formats for scanning video fields, synchronous multiplexing schemes, and synchronizing and framing techniques employed in data transmission. In fact, in all forms of data transmission, it seems that some form of periodicity is imposed on the signal format. Random processes with statistical properties that vary periodically with time are encountered frequently, not only in electrical communication systems, but in biological systems, chemical processes, and studies concerned with meteorology, ecology, and other physical and natural sciences. Systems analysts have tended, for the most part, to treat these "cyclostationary" processes as though they were stationary. This is done simply by averaging the statistical parameters (mean, variance, etc.) over one cycle. This averaging is equivalent to modelling the time-reference or phase of the process as a random variable uniformly distributed over one cycle. This type of analysis is appropriate in situations where the process is not observed in synchronism with its periodic structure. However, in a receiver that is intended for a cyclostationary signal, there is usually provided a great deal of information--in the form of a synchronizing pulse-stream or a sinusoidal timing signal--about the exact phase of the signal format. Most systems are, in fact, inoperative without this information.			

ACKNOWLEDGEMENT

I would like to express my gratitude to Professor L.E. Franks for suggesting to me the area of research to which this report is devoted. I would also like to acknowledge partial support of the research reported herein by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-71-2111.

ABSTRACT

Random signal processes which have been subjected to some form of repetitive operation such as sampling, scanning or multiplexing will usually exhibit statistical properties which vary periodically with time. In many cases, the repetitive operation is introduced intentionally to put the signal in a format which is easily manipulated and which preserves the time-position integrity of the events which the signal is representing. Familiar examples are radar antenna scanning patterns, raster formats for scanning video fields, synchronous multiplexing schemes, and synchronizing and framing techniques employed in data transmission. In fact, in all forms of data transmission, it seems that some form of periodicity is imposed on the signal format.

Random processes with statistical properties that vary periodically with time are encountered frequently, not only in electrical communication systems, but in biological systems, chemical processes, and studies concerned with meteorology, ecology, and other physical and natural sciences.

Systems analysts have tended, for the most part, to treat these "cyclostationary" processes as though they were stationary. This is done simply by averaging the statistical parameters (mean, variance, etc.) over one cycle. This averaging is equivalent to modelling the time-reference or phase of the process as a random variable uniformly distributed over one cycle. This type of analysis is appropriate in situations where the process is not observed in synchronism with its periodic structure. However, in a receiver that is intended for a cyclostationary signal, there is usually provided a great deal of information--in the form of a

synchronizing pulse-stream or a sinusoidal timing signal--about the exact phase of the signal format. Most systems are, in fact, inoperative without this information.

The first chapter of this dissertation is introductory and features a detailed historical account of the meager development and application of the theory--still in its infancy--of cyclostationary processes (as it appears in the engineering literature).

The second chapter is an extensive treatment of the topics of transformation, generation, and modelling of cyclostationary processes; and, among other things, serves to introduce a large number of models for cyclostationary processes. These models are used throughout the dissertation for illustrating various theoretical results.

The third chapter is an in-depth treatment of series representations for cyclostationary processes, and their autocorrelation functions, and other periodic kernels. These representations are applied to the problems of analysing cyclostationary processes, solving linear integral equations with periodic kernels, realizing periodically time-varying linear systems, and defining a generalized Fourier transform for cyclostationary (and stationary) processes.

The fourth chapter addresses itself to the problem of least-mean-squared-error linear estimation (optimum filtering) of cyclostationary processes, and employs the representations of Chapter III to obtain solutions, and the models of Chapter II to illustrate these solutions. Previous analyses of optimum filtering operations have assumed the stationary model for these processes and result in time-invariant filters. One of the major results in this thesis is the demonstration of the improvement in

performance that can be obtained by recognizing that--by virtue of timing information at the receiver--the received process is actually cyclostationary and the optimum filter is a periodically time-varying system. Numerous illustrative examples including amplitude-modulation, frequency-shift-keying, pulse-amplitude-modulation, frequency-division-multiplexing, and time-division-multiplexing are worked out in detail and include realizations of the optimum time-varying filters.

TABLE OF CONTENTS

Title Page	i
Acceptance Page	ii
Acknowledgement	iii
Abstract	iv
List of Figures	ix
I. INTRODUCTION	1
1. Motivation and Brief Description	1
2. Definition of Cyclostationarity	5
3. Historical Notes	12
4. Preliminary Comments on Mathematical Rigor	17
II. TRANSFORMATION, GENERATION, AND MODELING OF CYCLOSTATIONARY PROCESSES	25
1. Introduction	25
2. Linear Transformations	27
3. Multi-dimensional Linear Transformations (Scanning)	52
4. Nonlinear Transformations	57
5. Random Linear Transformations	66
6. Random Multi-dimensional Linear Transformations (Random Scanning)	91
7. Nonlinear Random Transformations (Jitter)	93

III. <u>SERIES REPRESENTATIONS FOR CYCLOSTATIONARY PROCESSES</u>	
(AND THEIR AUTOCORRELATION FUNCTIONS)	114
1. Introduction	114
2. Translation Series Representations	119
3. Harmonic Series Representations	165
4. Fourier Series Representation for Autocorrelation Functions	179
IV. LEAST-MEAN-SQUARED-ERROR LINEAR ESTIMATION OF CYCLOSTATIONARY PROCESSES	210
1. Introduction	210
2. The Orthogonality Condition	210
3. Optimum Time-Invariant Filters for Stationary Processes	226
4. Optimum Time-Invariant Filters for Cyclostationary Processes	230
5. Optimum Time-varying Filters for Cyclostationary Process	233
6. Examples of Optimum Filters and Improvements in Performance	266
V. SUMMARY	324
1. Summary of Chapters II, III, IV	325
2. Suggested Topics for Further Research	344
Bibliography	348

LIST OF FIGURES

Figure	Page
(1-1) Venn diagram for random processes	19
(1-2) Sampled and held process	20
(1-3) Autocorrelation function for a sampled and held process	21
(1-4) Time-division-multiplexed signal	22
(1-5) Line-scanning diagram	23
(1-6) Autocorrelation function for video signal	24
(2-1) System-Diagram	101
(2-2) Structural Diagram for Fourier Series Representation of periodically time-varying linear system	102
(2-3) Generation of PAM	103
(2-4) TDM multiplexor	104
(2-5a) FDM multiplexor	105
(2-5b) Diagram of support for double Fourier transform of autocorrelation function for FDM signal	106
(2-6) Time-scale compression transformation	107
(2-7) Time-scale compressed process	108
(2-8) Line-scanning diagram	109
(2-9) Line-scanning functions	110
(2-10) Circle-scanning diagram	111
(2-11) Circle-scanning function	112
(2-12) Window function	113

(3-1) Multiplier-integrator realization of resolution device	195
(3-2) Filter-sampler realization of resolution device	196
(3-3) Adaptation of multiplier-integrator realization	197
(3-4) Realization of resolution device for continuous computation	198
(3-5) TSR resolution and reconstruction of a T-CS process	199
(3-6) TSR resolution into real and imaginary parts	200
(3-7) TSR resolution and reconstruction of a T-CS process	201
(3-8) Autocorrelation-function shape	202
(3-9) Implementation of a periodically time-varying system	203
(3-10) Resolution of a process into its HSR representors	204
(3-11) HSR Resolution and reconstruction of a T-CS process	205
(3-12) Implementation of a periodically time-varying system	206
(3-13) Generation of processes whose crosscorrelations are the FSR coefficient-functions	207
(3-14a) Implementation of a periodically time-varying system	208
(3-14b) Implementation of a periodically time-varying system	209
(4-1) Communication-system diagram	301
(4-2) Orthogonal projection	302
(4-3) Structure of an M-dimensional optimum filter	303
(4-4a) Structure of optimum filter derived using TSR	304
(4-4b) Alternate structure of optimum filter derived using TSR	305
(4-4c) Structure of optimum filter when TSR correlation matrix is diagonal	306

(4-5) Structure of optimum filter derived using HSR	307
(4-6) Structure of optimum fixed-interval smoother	308
(4-7) Structure of optimum filter derived using FSR	309
(4-8) System diagram of optimum filter	310
(4-9) Feedback configuration for optimum filter	311
(4-10) Realization of operators for example 1	312
(4-11) Realization of operators for example 2	313
(4-12) Realization of operators for example 3	314
(4-13) Realization of operators for example 4	315
(4-14) Structure of optimum time-varying filter for FSK signal	316
(4-15) Structure of optimum time-varying filter for TDM-PAM signal	317
(4-16) Structure of optimum time-varying filter for FDM signal	318
(4-17) Estimation error for video signal in additive noise	319
(4-18) Structure of optimum time-varying filter for FDM signal	320
(4-19) Optimum time-varying filter for periodically attenuated signal in additive noise	321
(4-20) Gating functions	322
(4-21) Optimum time-varying filter for periodically time-scaled signal in additive noise	323

CHAPTER I

INTRODUCTION

1. Motivation and Brief Description

Random signal processes which have been subjected to some form of repetitive operation such as sampling or scanning will usually exhibit statistical properties which vary periodically with time. In many cases, the repetitive operation is introduced intentionally to put the signal in a format which is easily manipulated and which preserves the time-position integrity of the events which the signal is representing. Familiar examples are radar antenna scanning patterns, raster formats for scanning video fields, synchronous multiplexing schemes, and synchronizing and framing techniques employed in data transmission. In fact, in all forms of data transmission, it seems that some form of periodicity is imposed on the signal format.

Random processes with statistical properties that vary periodically with time are encountered frequently, not only in electrical communication systems, but in studies concerned with biological systems, meteorology, chemical processes, and ecological systems.

One process which exhibits statistically periodic characteristics is a synchronously timed pulse-sequence as used in digital data-transmission. Bennett [1] recognized this as a special case of a broader class of processes which he termed cyclostationary [2]. Other authors [3]-[9] have used the terms "periodically stationary", "periodic-stationary", "periodically correlated", "periodic-nonstationary" and "periodically nonstationary" to denote this same class.

Systems analysts have tended, for the most part, to treat these processes statistically as though they were stationary processes. This is done simply by averaging the statistical parameters (mean, covariance, etc.) over one period. This averaging, is equivalent to assuming that the time-reference, or phase, of the repetitive operation is completely indeterminate; i.e., the phase is a random variable, uniformly distributed over one period. This type of analysis is appropriate in situations where the process is not observed (or measured) in synchronism with its periodic structure. For example, a cyclostationary process may be an interference in a signal-transmission-channel whose receiver has no knowledge of the phase of the interfering process. In such a situation, the concepts used with stationary processes, such as power spectral density, can be valuable in evaluating system performance. However, in a receiver that is intended for a cyclostationary process, there is usually provided a great deal of information--in the form of a synchronizing pulse-stream or a sinusoidal timing signal--about the exact phase of the signal format. Most systems are, in fact, inoperative without this information.

Previous analyses of optimum-receiver filtering operations have assumed the stationary model for these processes and--for least-mean-squared-error reconstruction of the signal--result in time-invariant filters. One of the major results in this thesis is the demonstration of the improvement in performance that can be obtained by recognizing that--by virtue of timing information at the receiver--the received process is actually cyclostationary and the optimum receiver-filter is

a periodically time-varying system. Numerous illustrative examples including amplitude-modulation, frequency-shift-keying, pulse-amplitude-modulation, frequency-division-multiplexing, and time-division-multiplexing are worked out in detail and include realizations of the optimum time-varying filters.

The many examples presented not only demonstrate improvement, but also illustrate the general methods which are developed for representing cyclostationary processes and solving for and synthesizing optimum periodic filters. These methods are embodied in a general approach which applies not only to the problems of representation and estimation of cyclostationary processes, but more extensively to the solution of linear integral equations with periodic kernels and the synthesis of periodic linear systems.

The remainder of this introductory chapter serves to define cyclostationarity, and to illustrate same with several examples of cyclostationary signals which are commonly employed in communication systems, and to give a historical account of the development and application of the theory of cyclostationary processes (as it appears in the engineering literature).

The second chapter is an extensive treatment of the topics of transformation, generation, and modeling of cyclostationary processes; and, among other things, serves to introduce a large number of classes, or types, of models for cyclostationary processes. These models are used throughout the dissertation for illustrating various theoretical results.

The third chapter is an in-depth treatment of series representations for cyclostationary processes, and their autocorrelation

functions, and other periodic kernels; and includes brief discussions of their application to the solution of integral equations, and to the realization of periodically time-varying systems.

The fourth chapter addresses itself to the problem of least-mean-square linear estimation of cyclostationary processes (in noise), and employs the representations of Chapter III to obtain solutions, and the models of Chapter II to illustrate these solutions.

The fifth chapter is a detailed summary of the contents of Chapters II, III, IV and includes recommendations for further research. As a convenience for the reader, this final chapter serves as a combination index, outline and summary. It is hoped that this detailed summary will help to offset the length of Chapters II, III, IV.

2. Definition of Cyclostationarity

We begin with definitions for two types of stationarity:

DEFINITION: A random process $\{x(t); t \in J\}$ is stationary-of-order-two (S(2)) if and only if:

- i) the probability density functions (PDF's) for the random variables $\{x(t)\}$ are identical for all $t \in J$:

$$P_{x(t)}(\cdot) = P_{x(0)}(\cdot) \quad \forall t \in J,$$

- ii) for every $t, s \in J$ the joint PDF for the random variables $x(t)$ and $x(s)$ depends only on the difference of indices $t-s$:

$$P_{x(t)x(s)}(\cdot, \cdot) = P_{x(t-s)x(0)}(\cdot, \cdot) \quad \forall t, s \in J.$$

If x is a continuous-time random process, then the index set J is the set of real numbers R , and if x is a random sequence¹ then J is the set of integers I .

1. We are using an extended definition of a random sequence; strictly speaking, the domain of a sequence is the natural numbers (positive integers). Notice also that if x is a random sequence, then the PDF's $P_{x(0)}(\cdot)$, $P_{x(t)x(0)}(\cdot, \cdot)$ are composed of one- and two-dimensional impulse functions.

DEFINITION: A random process $\{x(t); t \in J\}$ is stationary-in-the-wide-sense (WSS) if and only if:

i) the mean function for the random process x is independent of the index t for every $t \in J$:

$$m_x(t) = E\{x(t)\} = \int_{-\infty}^{\infty} \sigma p_{x(t)}(\sigma) d\sigma = m_x(0) \quad \forall t \in J,$$

ii) for every $t, s \in J$ the autocorrelation function for the random process x depends only on the difference of indices $t-s$:

$$\begin{aligned} k_{xx}(t, s) &= E\{x(t)x^*(s)\} = \iint_{-\infty}^{\infty} \sigma \gamma^* p_{x(t)x(s)}(\sigma, \gamma) d\sigma d\gamma \\ &= k_{xx}(t-s, 0) \quad \forall t, s \in J. \end{aligned}$$

In the above definition, $E\{\cdot\}$ denotes statistical expectation, and $*$ denotes complex conjugation. In the interest of notational economy, we will (with some abuse of notation) identify the autocorrelation function for a WSS process x as $k_{xx}(t-s) = k_{xx}(t-s, 0)$.

Notice that stationary-of-order-two processes compose a subclass of the class of processes which are wide-sense stationary, and both of these classes are characterized by the invariance of certain statistical parameters to arbitrary shifts of the indexing variable. We now define two superclasses which are characterized by the invariance of certain statistical parameters to shifts of the indexing variable by multiples of a basic period. To emphasize the cyclic character of these processes, the term cyclostationary will be used:

DEFINITION: A random process $\{x(t); t \in J\}$ is cyclostationary-of-order-two (CS(2)) with period T if and only if:

i) the PDF for the random variable $x(t)$ is identical to that for the

random variable $x(t+T)$ for every $t \in J$:

$$P_x(t)(\cdot) = P_x(t+T)(\cdot) \quad \forall t \in J,$$

ii) for every $t, s \in J$ the joint PDF for the random variables $x(t)$ and $x(s)$ is identical to that for the random variables $x(t+T)$ and $x(s+T)$:

$$P_{x(t)x(s)}(\cdot, \cdot) = P_{x(t+T)x(s+T)}(\cdot, \cdot) \quad \forall t, s \in J.$$

DEFINITION: A random process $\{x(t); t \in J\}$ is cyclostationary-in-the wide-sense (WSCS) with period T if and only if:

i) the mean function for the random process x is T -periodic:

$$m_x(t+T) = E\{x(t+T)\} = m_x(t) \quad \forall t \in J,$$

ii) for every $t, s \in J$ the autocorrelation function for the random process x is jointly T -periodic in t and s :

$$k_{xx}(t+T, s+T) = E\{x(t+T)x^*(s+T)\} = k_{xx}(t, s) \quad \forall t, s \in J.$$

Notice that condition ii) in this last definition--that $k_{xx}(\cdot, \cdot)$ be jointly T -periodic in its two variables--is less restrictive than the condition that $k_{xx}(\cdot, \cdot)$ be individually T -periodic in each of its variables:

$$k_{xx}(t+T, s) = k_{xx}(t, s+T) = k_{xx}(t, s) \quad \forall t, s \in J,$$

which if substituted for ii) would define the class of processes which are periodic-in-the-wide-sense (WSP).

Notice also that it is possible--in fact, occurs frequently in practice--for WSCS processes to have constant mean and variance ($k_{xx}(t, t) = m_x^2(t)$) like WSS processes, and yet have autocorrelation functions which display periodic fluctuations.

Now, all processes which are stationary-of-order-two are cyclostationary-of-order-two with any choice of period T , and all wide-sense stationary processes are wide-sense cyclostationary with any choice of period T , and all wide-sense periodic processes are wide-sense cyclostationary and are not² wide-sense stationary. In summary, these various classes of processes are related as shown in the Venn diagram of Figure (1-1).

We will be primarily concerned with WSCS processes in this thesis and, from this point on, will refer to these simply as CS.

Most of the cyclostationary processes studied in this thesis are derived from stationary processes which have been subjected to some form of repetitive operation such as sampling, scanning, multiplexing, or time-scale modulation. There is, in fact, a large variety of mechanisms by which periodicity is imposed on a process which might otherwise be stationary. Following are several representative examples:

i) Sampling and synchronous data signals. A particularly interesting example of a cyclostationary process is the signal x which results from synchronously sampling a WSS process y and reconstructing the waveform with a prescribed interpolating pulse. Realizations of this process take the form:

$$x(t) = \sum_{n=-\infty}^{\infty} y(nT)q(t-nT) \quad (1-1)$$

where $\{y(nT)\}$ are the sample values of the realization $y(\cdot)$, and $q(\cdot)$ is the interpolating pulse. For example, the familiar sample-and-hold

²This is so except for the single degenerate case where $k_{xx}(t,s) = \text{constant}$ for all $t,s \in J$.

operation is characterized by a rectangular $q(\cdot)$ of unit height and width T . A typical realization of a sampled-and-held process is shown in Fig. (1-2). The mean and autocorrelation for the process are given by the expressions [2]:

$$\begin{aligned} m_x(t) &= m_y \sum_n q(t-nT) \\ k_{xx}(t,s) &= \sum_{n,m} k_{yy}(nT-mT) q(t-nT) q(s-mT) \end{aligned} \quad (1-2)$$

and cyclostationarity of x is easily verified. Notice that, for the sample-and-hold pulse, x has constant mean and variance but its autocorrelation function has strong periodic fluctuations which are shown graphically in Figure (1-3). (The periodic fluctuations of a CS process are displayed along lines parallel to the $(t = s)$ diagonal.)

Another version of this process is synchronous pulse-amplitude-modulation (PAM) where the numbers $\{y(nT)\} = \{y_n\}$ are realizations of a random data-sequence, and the pulse shape $q(\cdot)$ is selected to minimize the ill effects of intersymbol-interference [2] which result from transmission over a dispersive communication-channel. This model can be made more realistic by modifying the time-position of each pulse with a random jitter variable:

$$x(t) = \sum_n y_n q(t-nT+\delta_n).$$

The resultant process termed jittered PAM can still be cyclostationary with period T [2] (See Sec. 7 of Chapter II).

Other pulse-modulation techniques such as frequency-shift-keying, phase-shift-keying, pulse-width-modulation, and pulse-position-modulation also give rise to synchronous pulse-trains which can be modeled as CS processes [10] (see Sec. 4 of Chapter II). In fact even

asynchronous pulse-trains can be modeled as CS processes as shown in Sec. 2 of Chapter II.

ii) Scanning and time-division-multiplexed signals. In communications, one often encounters processes formed by interleaving finite-length records from a multiplicity of signal sources. As an example consider the process with realizations of the form

$$x(t) = \sum_{n=-\infty}^{\infty} x_n(t)q(t-nT) \quad (1-3)$$

where $q(t-nT)$ is a "gating pulse" (evaluated at time t) which "turns on" the n^{th} signal source in the time-interval $[nT, (n+1)T]$, and $\{x_n\}$ are the various random signals. Specifically, consider the time-division-multiplex (TDM) of M different jointly WSS signals $\{x_n; x_n = x_{n+M} \forall n\}$ with crosscorrelations $k_{n\ell}(t-s) = E\{x_n(t)x_\ell(s)\}$ and means $m_n = E\{x_n(t)\}$. A typical realization of such a TDM signal is shown in Fig. (1-4). The mean and autocorrelation for the composite TDM process are given by the expressions:

$$\begin{aligned} m_x(t) &= \sum_n m_n q(t-nT) \\ k_{xx}(t,s) &= \sum_{n,\ell} k_{n\ell}(t-s)q(t-nT)q(s-\ell T) \end{aligned} \quad (1-4)$$

and the cyclostationarity (period MT) is easily verified. Notice that regardless of the differences among the various auto- and cross-correlation functions for the M signals, if the means and variances are equal, the mean and variance of the composite cyclostationary process are both constant (for a rectangular gate function). Yet its autocorrelation function can have pronounced periodic variations even when all M of the signals are statistically identical (in which case the period

of cyclostationarity is T). Other CS models for TDM signals are derived in Sec. 2 of Chapter II.

As another special case of those processes which take the form of Equation (1-3), consider the video signal which results from conventional line-scanning of a rectangular visual-field [11]. Here the $\{x_n\}$ represent consecutive one-line segments of the video signal and M is the number of lines per frame. Using the Markov model derived by Franks [11] for the line segments (see Figure (1-5)) we obtain

$$k_{nl}(t-s) = \rho^{|n-l|} e^{-2\pi f_0 |t-s| \bmod T}, \quad (t-nT), (s-lT) \in [0, T].$$

Thus, the video autocorrelation function becomes (from Equation (1-4)):

$$k_{xx}(t,s) = e^{-2\pi f_0 |t-s| \bmod T} \sum_{n,l} \rho^{|n-l|} q(t-nT)q(s-lT) \quad (1-5)$$

which, interestingly, is the autocorrelation function of a sampled and held process (Equation (1-2)) modified by a factor--call it $k_1(t-s)$ --which is T -periodic in $t-s$. This periodic factor is shown graphically in Fig. (1-6), where the parameter f_0 has been chosen much greater than $1/T$ as it is for typical viewing material. See Secs. 3,6 of Chapter II for a more detailed model of the video process which takes into account frame-to-frame correlation as well as line-to-line correlation.

Many other CS processes such as amplitude-modulated signals, analog and digital phase and frequency modulated signals, frequency-division-multiplexed signals, etc. are introduced and discussed in Chapter II.

3. Historical Notes

Interest in cyclostationary random processes seems to have originated in the investigations of a group of communications engineers and scientists involved in the development of pulse-code-modulation (PCM) at Bell Telephone Laboratories³ in the late 1940's [4]. However, the cyclostationarity of these PCM processes did not specifically begin to receive attention until the late 1950's. In 1958, Bennett [1] developed a statistical model for the synchronous pulse-train of PCM and demonstrated that the process is nonstationary and exhibits periodically varying mean and autocovariance functions. He coined the term "cyclostationary" to denote the class of nonstationary processes which exhibit these cyclic characteristics, and he introduced the Fourier series representation for the mean and autocovariance functions. In addition, Bennett discussed the phase-randomization method of obtaining a stationary model for a cyclostationary process, and he pointed out that the power spectral density obtained from this model is identical to that obtained by Macfarlane [12] in 1949 using Rice's general definition of the average power spectral density of a nonstationary process [13].

Four years prior to Bennett's publication, Deutsch [8] presented a paper on the envelope detection of cyclostationary processes--which he referred to as "periodically-stationary"--composed of one or more periodic signals each amplitude modulated with a stationary Gaussian process. Although this article preceded Bennett's, it did not deal primarily with the periodic structure of the cyclostationary processes.

³ The basic ideas of--and original patents on--PCM date back to the early 1930's. See Techniques of Pulse Code Modulation for a historical review [39].

At about the same time (late 1950's), interest in cyclostationary processes was initiated in Russia with a series of theoretical works on radiophysics. Gudzenko, motivated by these works, published a paper in 1959 on the general structure of cyclostationary processes which he referred to as "periodically nonstationary" [14]. He briefly examined the periodic correlation function and corresponding "time-varying spectrum" for the general (arbitrary) cyclostationary process, and gave sufficient conditions for the consistent estimation of the coefficient functions in the Fourier series expansion of the autocorrelation function. He also mentioned the interpretation of the zeroth order Fourier coefficient as the autocorrelation function of a stationary process, and briefly discussed linear filtering of cyclostationary processes. Following Gudzenko's paper were two works published by Gladyshev who was associated with the Institute of Atmospheric Physics in the Academy of Sciences of the USSR [15,16]. The first paper (1961) was an investigation of the mathematical structure of cyclostationary sequences, and the second (1963) was a more brief investigation of the properties of continuous-time cyclostationary and "almost cyclostationary" processes. Specifically, he considered the non-negative definiteness of the Fourier series coefficient functions, and the harmonizability of these processes which he referred to as "periodically and almost periodically correlated".

Following these pioneering works, there have been a number of publications dealing with cyclostationary processes as related to meteorology, signals for communication, and noise in electrical devices and circuits. Jordan (1961) derived the least-squares series representation,

on a finite interval, of a random signal in the presence of noise, and briefly discussed its application to the problem of least-squares estimation of cyclostationary signals in noise [25]. Monin (1963) suggested using cyclostationary processes as models for meteorological processes [17]. Willis (1964), motivated by a study of cosmic ray extensive air showers, investigated a non-homogeneous Poisson process with periodic rate parameter, and indicated how the amplitude and frequency of the periodic rate-of-occurrence of events can be estimated from measured time intervals between events [18]. Parzen (1962) suggested that a non-homogeneous Poisson process with periodic rate parameter would serve as a suitable model of electron emissions from the cathode of a temperature-limited diode with alternating filament-current [22]. Markelov (1966) studied some statistical properties of the cyclostationary output of a parametric amplifier driven by a stationary process. He derived the average number of axis-crossings and relative staying time in an interval [21]. Anderson and Salz (1965) derived the power spectral density for the stationarized version of a cyclostationary digital FM signal [23], and Lundquist (1969) similarly derived the power spectral density for digital PM [24].

Several additional references to work related to cyclostationary processes can be found in the references in Brelsford's doctoral dissertation "Probability Predictions and Time Series with Periodic Structure", (1967) [19]. Chapter V of his dissertation is devoted to a brief review of representations for cyclostationary processes (mostly sequences) and the presentation of autoregressive techniques for the

linear prediction of cyclostationary sequences. In particular, Breilsford (and Jones [20]) demonstrated--using meteorological data--that autoregressive techniques which employ periodically varying coefficients can yield better predictions of cyclostationary sequences than prior techniques which used only constant coefficients.

A few other references to work related to cyclostationary processes can be found in the bibliography and historical notes in Hurd's doctoral dissertation "An Investigation of Periodically Correlated Stochastic Processes" (1965) [5]. In his dissertation, Hurd investigates (with mathematical rigor) the structure and experimental analysis of cyclostationary processes which he refers to as "periodically correlated". His work centers around the Fourier series representation for the correlation functions for cyclostationary processes, and is closely related to the pioneering works of Gudzenko (1959) and Gladyshev (1963). Hurd's dissertation is divided into two main areas. The first contains an examination of the general mathematical structure of cyclostationary processes and their correlation functions, and contains a brief discussion of the stationarizing effect of a random phase and the effects of linear filtering, and includes a few models for the generation of cyclostationary processes from stationary processes. The second area contains treatments of several topics in experimental analysis. Specifically, cyclostationary processes are shown to have "ergodic" properties in the sense that the autocorrelation function and its Fourier series coefficient-functions can--under certain conditions-- be estimated from one sample function of a cyclostationary process. Also, some methods for testing observed waveforms for cyclostationarity are presented along with some experimental results.

The textbooks of Stratonovich (1963) [7], Papoulis (1965) [3], Bendat and Piersol (1966) [9], and Franks (1969) [2], define the class of cyclostationary processes, give various examples, and (except [9]) discuss the stationarizing effect of phase-randomization. The most extensive treatment is that of Franks. In a chapter devoted mostly to the development of models for cyclostationary processes, he points out that most scanning operations, such as radar antenna circular-scanning and video line-scanning, can produce cyclostationary signals, and in fact that many of the commonly used repetitive signal processing operations including sampling, scanning, and multiplexing can transform stationary processes into cyclostationary processes. He develops, in detail, a cyclostationary model for synchronous pulse-amplitude-modulation (PAM) and shows that "jittered" PAM (non-synchronous) can also be cyclostationary. He also introduces cyclostationary models for the time-division multiplex of two stationary signals, and for synchronous pulse-width-modulation.

Finally, we mention the very recent work of Ogura (1971) in which he briefly discusses series and integral representations for cyclostationary processes, and presents in some detail an interesting harmonic series representation for cyclostationary processes and their autocorrelation functions. He also briefly discusses the measurement of the coefficients in his harmonic series representation for "ergodic" cyclostationary processes, and concludes with three examples of cyclostationary processes: AM, PAM, and PPM.

In summary, the major works which deal directly with the general class of cyclostationary processes are, in chronological order, those of: Bennett (1958), Gudzenko (1959), Gladyshev (1961, 1963), Brelsford (1967), Franks (1969), Hurd (1969), and Ogura (1971).

4. Preliminary Comments on Mathematical Rigor

Although most engineers are familiar with and use the Riemann integral [26] rather than the Lebesgue integral [27], it is the latter which is indispensable in analyses where questions of the existence of integral and series representations of functions (random and deterministic) arise, and where the interchange of the order of execution of integration, infinite summation, and statistical expectation must be justified. (See, for example, Wiener [28] pp. 4-5, and Epstein [29] p. 29). Such questions of existence and needs for justification arise throughout every remaining chapter of this dissertation, and can only be rigorously dealt with in terms of Lebesgue integration theory and associated measure-theoretic notions. For example, the work of Hurd [5], which is closely related to this thesis, is thoroughly based in measure theory.

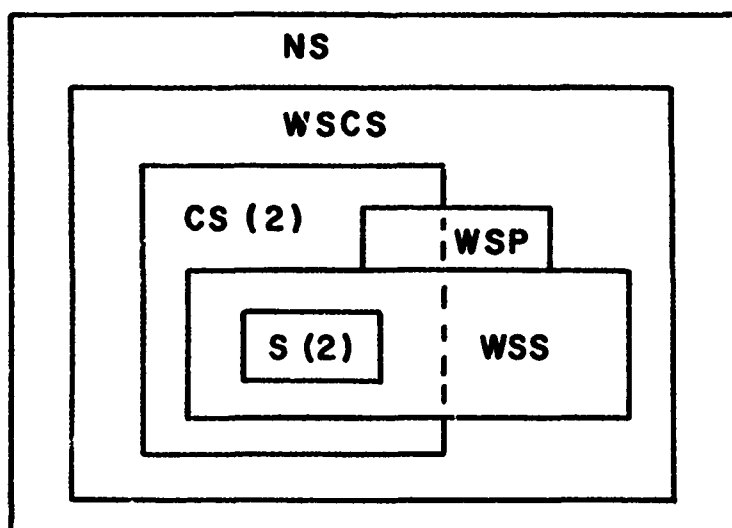
Despite these comments, it is not my intent to clutter the presentation of the methods, techniques, and results of this thesis with rigorous justifications of the numerous manipulations, and with precise statements of the conditions under which the various (known) representations are valid. Such an undertaking would surely transform this engineering thesis into a mathematical work of limited appeal to most engineers.

Rather, I will assume that all random processes of interest are harmonizable in the sense that the double Fourier transform of their autocorrelation functions exist (with the possible inclusion of impulse functions); that all periodic functions (including periodic trains of impulses) possess Fourier series representations; and that wherever necessary, the order of execution of integration, infinite summation, and statistical expectation can be interchanged. Clearly then, all proofs based on these assumptions will be formal proofs.

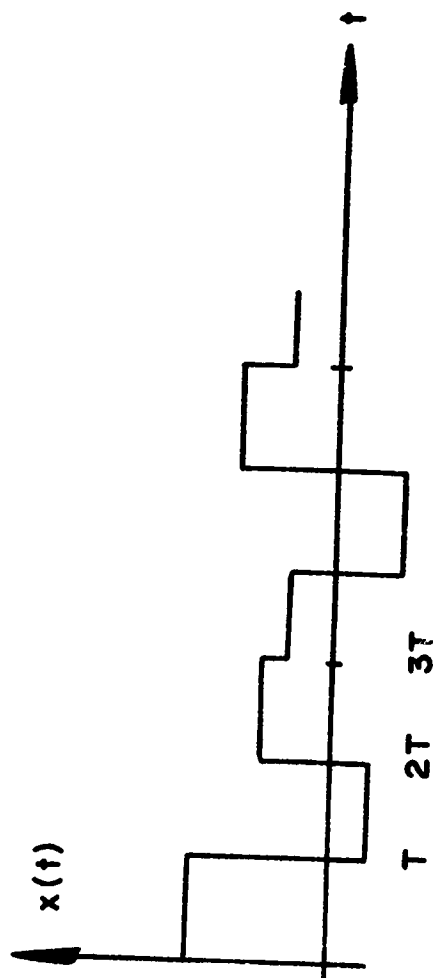
I refer the mathematically inclined reader, interested in the validity of these assumptions, to the following references chosen on the basis of their potential appeal to engineers:

- (1) Lebesgue integration theory: Riesz and Sz-Nagy [31], and Epstein [29],
- (2) Properties of the Lebesgue integral, and development of series and integral harmonic representations for deterministic functions: Wiener [28], and Titchmarsh [32,33],
- (3) Theory of linear integral equations: Riesz and Sz-Nagy [31],
- (4) Theory of Distributions (impulse functions and their relatives): Zemanian [30],
- (5) Development of series and integral representations for random functions (introductory treatments for engineers): Pugachev [34], and Papoulis [3],
- (6) Linear functional analysis: Epstein [29], and Riesz and Sz-Nagy [31].

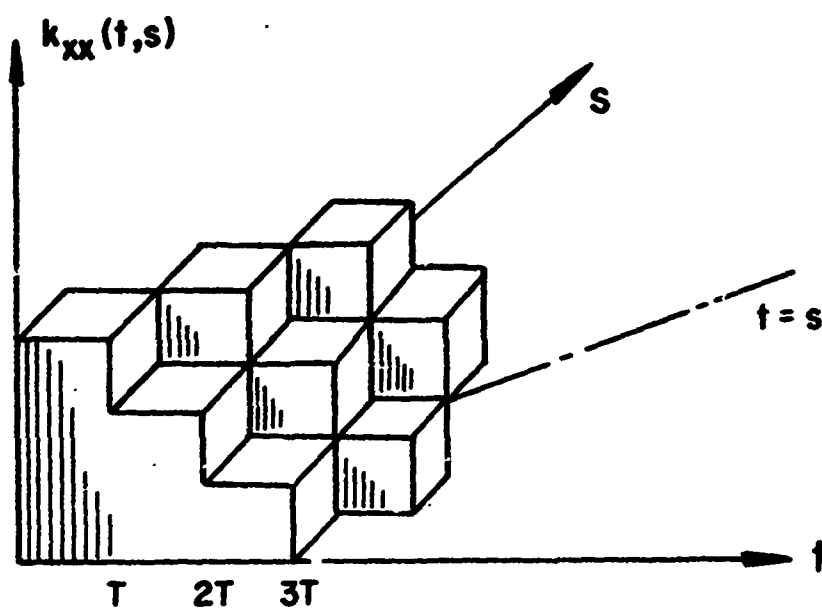
Finally, I mention that the use of the Fourier transforms and Fourier series (except in Chapter III, Section 2h where a generalized Fourier transform for random processes is defined) and the use and manipulation of infinite integrals, infinite series, and impulse functions in this dissertation are not uncommon in the engineering literature, and are often easy to justify heuristically.



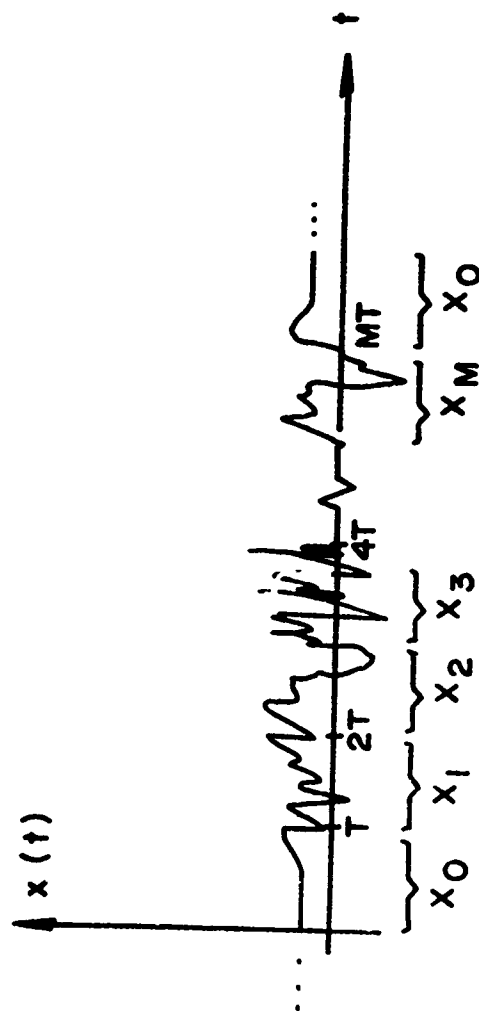
Figure(1-1) Venn diagram of various classes of random processes:
NS = generally nonstationary, WSCS = wide-sense cyclostationary,
WSS = wide-sense stationary, CS(2) = cyclostationary-of-order-
two, S(2) = stationary-of-order-two, WSP = wide-sense periodic.



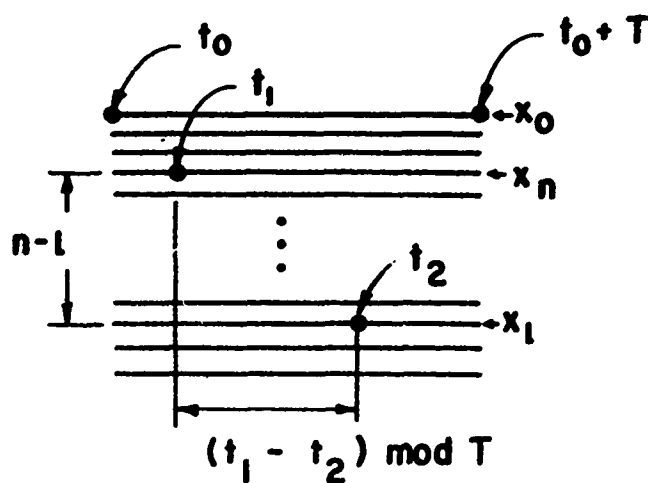
Figure(1-2) Realization of a sampled and held process.



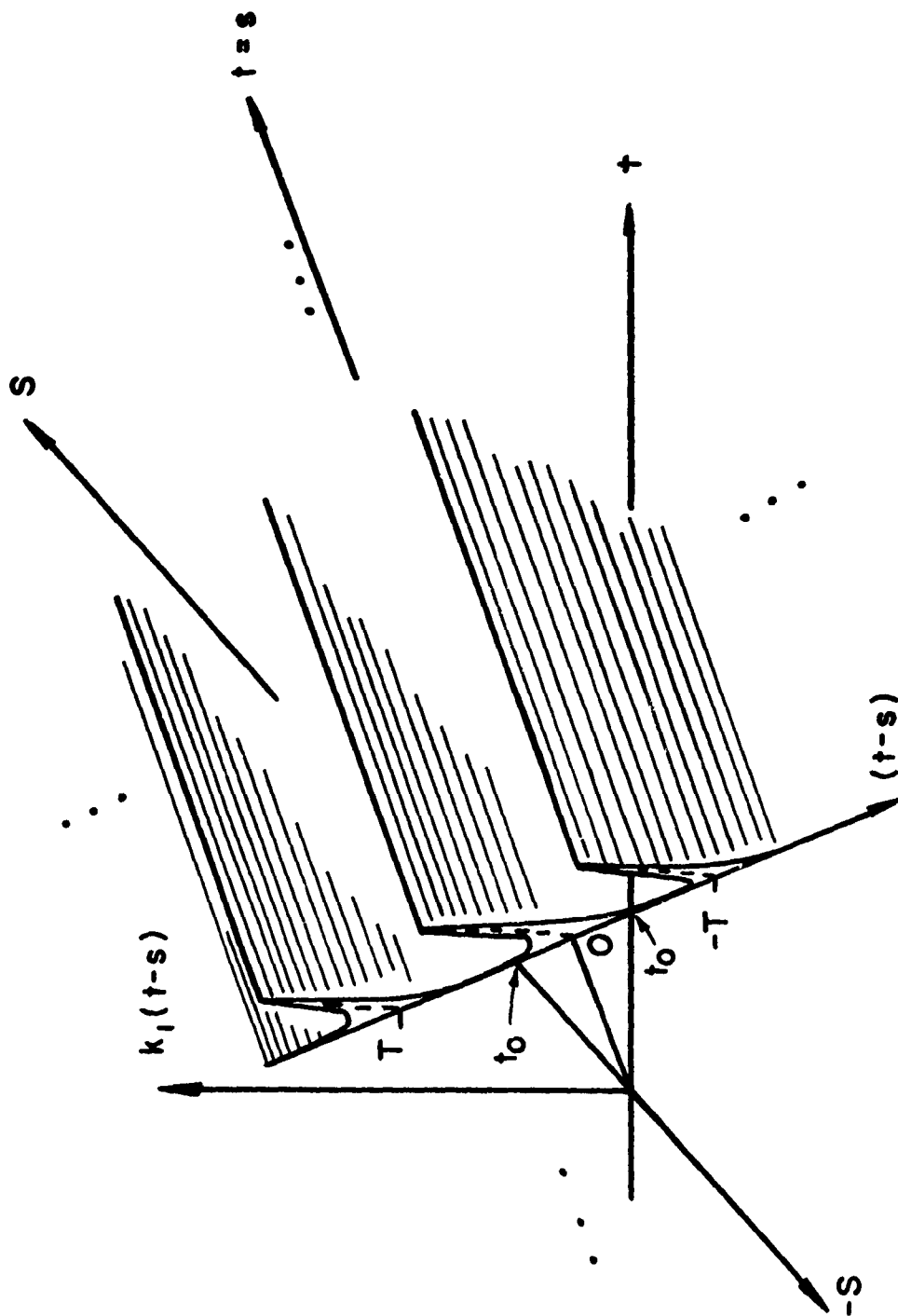
Figure(1-3) The autocorrelation function k_{xx} for a sampled and held process.



Figure(1-4) Realization of a time-division-multiplexed signal.



Figure(1-5) Line-scanning diagram.



Figure(1-6) Periodic factor in autocorrelation function for video signal.

CHAPTER II

TRANSFORMATION, GENERATION, AND MODELING OF CYCLOSTATIONARY PROCESSES

1. Introduction

This chapter provides a foundation upon which many of the developments in this thesis are based, and upon which subsequent studies dealing with cyclostationary processes can be based. We consider here the generation of cyclostationary processes and transformations on them. We develop formulas which describe, in terms of input-output relations, various classes of transformations such as linear time-invariant, periodically varying, and time-scale transformations; and random linear stationary and cyclostationary transformations. Also considered are nonlinear random and deterministic transformations encountered in various modulating schemes and common models of distortion. We derive models for a multitude of cyclostationary processes which are generated from stationary processes that have been subjected to transformations from the above classes. We also present a number of theorems on transformations mostly related to the generation and preservation of cyclostationarity.

Most of the developments in this chapter relate to the system diagram shown in Figure 2-1 where y is a random process referred to as the input to the system G , and x is a random process termed the output. x is also referred to as the image of y under the transformation G , and y is termed the inverse-image of x .

In Section 2, we consider Transformations G which are linear. After deriving general input-output relations, we individually discuss, in some detail, the three subclasses of linear transformations referred to as time-invariant filters (2a), periodically time-varying systems (2b), and time-scale transformations (2c). Zadeh's "system function" is introduced as a means for characterizing linear systems, and is used throughout most of this section. A number of theorems on the properties of these three classes of transformations are proven and illustrated with examples which also serve to introduce seven types or classes of cyclostationary processes including time-division-multiplexed signals, frequency-division-multiplexed signals, and a-synchronous pulse trains including facsimile, and telegraph signals.

In Section 3, we briefly discuss multidimensional linear transformations as a means for characterizing scanning operations. Examples are given to illustrate two common types of scanning, and to introduce two additional classes of cyclostationary processes including video signals.

In the fourth section, we introduce a generalization of Wiener's "Volterra series representation" for characterizing input-output relations for periodic nonlinear systems. Two subclasses of periodic Volterra systems are considered, and theorems for each, on the generation of cyclostationary processes, are proven and illustrated with examples which also serve to introduce several additional classes of cyclostationary processes, including synchronous pulse-trains such as frequency-shift-keyed signals.

Section 5, on random linear transformations, parallels Section 2 on deterministic linear systems. After deriving general input-output relations in terms of a generalization of Zadeh's "system correlation function", we individually discuss, in some detail, the three subclasses referred to as wide-sense-stationary systems (5a), cyclostationary systems (5b), and random time-scale transformations (5c). A number of theorems on the preservation and generation of cyclostationary processes are proven and illustrated with examples which also serve to introduce another eight classes or types of cyclostationary processes, including analog and digital frequency-modulated and phase-modulated signals.

Section 6, on random multidimensional linear transformations, parallels Section 3. The major application of the theory in this section is to random scanning.

The final section (7), on random nonlinear transformations, is an extension of Section 4 on deterministic nonlinear systems. The major issue in this section is "jitter". This random disturbance is characterized in general terms, and shown to preserve cyclostationarity under a fairly liberal set of conditions.

2. Linear Transformations

Due to the length of this section, we begin by giving a brief outline of the theorems proved and the cyclostationary process models defined:

a) Time-invariant filters

Theorem(2-1): Preservation of cyclostationarity by time-invariant filters.

Theorem(2-2): Sampling theorem for nonstationary processes.

Theorem(2-3): Reduction of cyclostationary processes to stationary processes via bandlimiting filters.

b) Periodically time-varying systems

Theorem(2-4): Generation of cyclostationary processes from stationary processes via periodic systems.

Model(1): Amplitude-modulated signals

Model(2): Pulse-amplitude-modulated signals

Model(3): Time-division-multiplexed signals

Model(4): Frequency-division-multiplexed signals

c) Time-scale transformations

Theorem(2-5): Generation of cyclostationary processes from stationary processes via periodic time-scale transformations.

Model(5): Time-division-multiplexed signals

Model(6): Doppler shifted processes

Model(7): Random facsimile signal

Model(8): Asynchronous pulse-amplitude modulated signals

We now turn to the development of general input-output relations which will be employed throughout the remainder of this thesis.⁴

⁴Throughout most of this dissertation, we will use single letters, say x , to denote random processes, and we will use the form $x(t)$ to denote a realization of x . Similarly, we will use the notation $y(t)$, $z(t,s)$ to denote deterministic functions of one and two variables. However, on occasion, the notation $y(\cdot)$, $z(\cdot, \cdot)$ will be used for the purpose of clarification. The index sets for random processes, and domains for deterministic functions will, unless otherwise stated, be $(-\infty, \infty)$ or cartesian products thereof.

All linear transformations of interest in this thesis are linear integral transformations and can be characterized by an impulse-response function (kernel) $g(t, \tau)$, which may possess impulse functions and their derivatives. Thus any realization $x(t)$ of the output of G can be related to its inverse-image (input) $y(t)$ as follows:⁵

$$x(t) = \int_{-\infty}^{\infty} g(t, \tau) y(\tau) d\tau. \quad (2-1)$$

The mean function for x is related to the mean function for y as follows:

$$\begin{aligned} m_x(t) \triangleq E\{x(t)\} &= E\left\{ \int_{-\infty}^{\infty} g(t, \tau) y(\tau) d\tau \right\} \\ &= \int_{-\infty}^{\infty} g(t, \tau) E\{y(\tau)\} d\tau \\ &= \int_{-\infty}^{\infty} g(t, \tau) m_y(\tau) d\tau, \end{aligned} \quad (2-2)$$

and the autocorrelation function for x is related to that for y as follows:

$$\begin{aligned} k_{xx}(t, s) \triangleq E\{x(t)x^*(s)\} &= E\left\{ \iint_{-\infty}^{\infty} g(t, \tau) g^*(s, \gamma) y(\tau) y^*(\gamma) d\tau d\gamma \right\} \\ &= \iint_{-\infty}^{\infty} g(t, \tau) g^*(s, \gamma) E\{y(\tau) y^*(\gamma)\} d\tau d\gamma \\ &= \iint_{-\infty}^{\infty} g(t, \tau) g^*(s, \gamma) k_{yy}(\tau, \gamma) d\tau d\gamma. \end{aligned} \quad (2-3)$$

The double Fourier transform of the autocorrelation function (sometimes referred to as the cointensity spectrum) is a quantity which finds much application in the theories of representation and estimation of random processes. We denote the double Fourier transform of $k_{xx}(t, s)$ as $K_{xx}(f, \nu)$:

⁵This characterization is generalized in Section 3 of this chapter to multidimensional transformations.

$$K_{xx}(f, v) \triangleq \iint_{-\infty}^{\infty} k_{xx}(t, s) e^{-j2\pi(ft-vs)} dt ds. \quad (2-4)$$

We can express the double Fourier transform of a function of two variables in terms of a repeated single Fourier transform as follows:

$$\text{let } \tilde{g}(t, \tau) \triangleq \int_{-\infty}^{\infty} g(t, \tau) e^{-j2\pi f\tau} d\tau;$$

$$\text{then } G(f, v) = \int_{-\infty}^{\infty} \tilde{G}(f, \tau) e^{j2\pi v\tau} d\tau.$$

With the aid of this definition of \tilde{G} we can easily relate $K_{xx}(f, v)$ directly to $K_{yy}(f, v)$ by combining Eqs. (2-3), (2-4):

$$\begin{aligned} K_{xx}(f, v) &= \iiint_{-\infty}^{\infty} k_{yy}(\tau, \gamma) g(t, \tau) g^*(s, \gamma) e^{-j2\pi(ft-vs)} d\tau d\gamma dt ds \\ &= \iint_{-\infty}^{\infty} k_{yy}(\tau, \gamma) \tilde{G}(f, \tau) \tilde{G}^*(v, \gamma) d\tau d\gamma. \end{aligned}$$

Now, using Parseval's relation [2] we have

$$\begin{aligned} K_{xx}(f, v) &= \iint_{-\infty}^{\infty} K_{yy}(\omega, \sigma) \left\{ \int_{-\infty}^{\infty} \tilde{G}^*(f, \tau) e^{-j2\pi\omega\tau} d\tau \right\}^* \cdot \\ &\quad \left\{ \int_{-\infty}^{\infty} \tilde{G}(v, \gamma) e^{j2\pi\sigma\gamma} d\gamma \right\} d\omega d\sigma \\ &= \iint_{-\infty}^{\infty} K_{yy}(\omega, \sigma) G(f, \omega) G^*(v, \sigma) d\omega d\sigma. \end{aligned} \quad (2-5)$$

Notice that Eq. (2-5) is the "frequency-time dual" of Eq. (2-3).

There are two interesting subclasses of linear transformations that are dealt with throughout this thesis. We describe these in the next two subsections.

a) Time invariant filters. The subclass, time-invariant linear transformations--frequently referred to as filters--is characterized by the defining relation (with some abuse of notation) $g(t, \tau) = g(t-\tau)$ for all t and τ . That is, the impulse-response function for a linear

time-invariant filter depends only on the difference of its arguments and is, in fact, a function of a single variable. The double Fourier transform of $g(t-\tau)$ exhibits the following singular behavior:

$$G(f, \nu) = G(f) \delta(f - \nu), \quad (2-6)$$

where $G(f)$ is the single Fourier transform of $g(t)$ and is referred to as the transfer function for the filter, and δ is the Dirac delta function (impulse function).

Using Eq. (2-6), the input-output relation of Eq. (2-5) reduces to

$$K_{xx}(f, \nu) = K_{yy}(f, \nu) G(f) G^*(\nu). \quad (2-7)$$

Now, if y is WSS then its autocorrelation function depends only on the difference of its arguments and, as a result,

$$K_{yy}(f, \nu) = K_{yy}(f) \delta(f - \nu) \quad (2-8)$$

where $K_{yy}(f)$ is the single Fourier transform of the autocorrelation function $k_{yy}(t)$ and is referred to as the power spectral density (PSD) for y (and sometimes as the autointensity spectrum). Using Eq. (2-8) in Eq. (2-7) yields

$$\begin{aligned} K_{xx}(f, \nu) &= K_{yy}(f) \delta(f - \nu) G(f) G^*(\nu) \\ &= K_{yy}(f) |G(f)|^2 \delta(f - \nu). \end{aligned}$$

Also, if y is WSS then its mean function is a constant and (from Eq. (2-2) with $g(t, \tau) = g(t - \tau)$) so too is the mean function for x . Thus, x is WSS and, from the last equation, has PSD

$$K_{xx}(f) = |G(f)|^2 K_{yy}(f). \quad (2-9)$$

The following theorem is an intuitively obvious, but useful, result on the preservation of the property of cyclostationarity by time-invariant filters:

THEOREM(2-1): If the input to an arbitrary time-invariant filter is a cyclostationary process with period T , then the output is cyclostationary with period T .

Proof: From Eq. (2-3) we have

$$\begin{aligned} k_{xx}(t+T, s+T) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t+T-\tau) g^*(s+T-\gamma) k_{yy}(\tau, \gamma) d\tau d\gamma \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t-\tau') g^*(s-\gamma') k_{yy}(\tau'+T, \gamma'+T) d\tau' d\gamma' \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t-\tau') g^*(s-\gamma') k_{yy}(\tau', \gamma') d\tau' d\gamma' \\ &= k_{xx}(t, s) \quad \forall t, s, \end{aligned}$$

and from Eq. (2-2)

$$\begin{aligned} m_x(t+T) &= \int_{-\infty}^{\infty} g(t+T-\tau) m_y(\tau) d\tau \\ &= \int_{-\infty}^{\infty} g(t-\tau') m_y(\tau'+T) d\tau' \\ &= \int_{-\infty}^{\infty} g(t-\tau') m_y(\tau') d\tau' \\ &= m_x(t) \quad \forall t. \end{aligned}$$

Hence, both the mean and autocorrelation for x are T -periodic so that x is T -CS.⁶

QED

⁶We will use the short-hand notation T -CS to denote cyclostationary (in the wide sense) with period T .

i) Bandlimiting.

(1) **Sampling Theorem:** The well known sampling theorem for WSS processes [2,3] may be stated as follows:

If the PSD $K_{xx}(f)$ of a WSS process x is bandlimited to the interval $(-1/2T, 1/2T)$ in the sense that $K_{xx}(f) = 0$ for $|f| \geq 1/2T$, then x admits the mean-square equivalent "sample representation"

$$E\left\{\left(x(t) - \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T}\right)^2\right\} = 0 \quad \forall t.$$

We now state and prove a generalization of this theorem which removes the restriction to WSS processes:

THEOREM(2-2): If the double Fourier transform $K_{xx}(f, \nu)$ of the autocorrelation function for a process x is bandlimited to $(-1/2T, 1/2T)$ in the sense that $K_{xx}(f, \nu) = 0$ if $|f| \geq 1/2T$ or if $|\nu| \geq 1/2T$, then x admits the mean-square equivalent representation

$$E\left\{\left(x(t) - \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T}\right)^2\right\} = 0 \quad \forall t.$$

Proof:

$$\text{Let } \epsilon \triangleq E\left\{\left(x(t) - \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T}\right)^2\right\}.$$

Expanding the square and interchanging expectation and summation yields

$$\begin{aligned} \epsilon &= k_{xx}(t, t) - 2 \sum_n k_{xx}(t, nT) \phi(t-nT) \\ &\quad + \sum_{n,m} k_{xx}(nT, mT) \phi(t-nT) \phi(t-mT) \end{aligned}$$

where

$$\phi(\tau) \triangleq \frac{\sin \pi \tau / T}{\pi \tau / T}.$$

Since $k_{xx}(t,s)$ is the inverse double Fourier transform of $K_{xx}(f,v)$, then

$$\begin{aligned} k_{xx}(t,s) &= \iint_{-\infty}^{\infty} K_{xx}(f,v) e^{j2\pi(ft-vs)} df dv \\ &= \iint_{-1/2T}^{1/2T} K_{xx}(f,v) e^{j2\pi(ft-vs)} df dv, \end{aligned}$$

and the last equality is a result of the bandlimiting hypothesis.

Substituting this last equation into the last expression for ϵ we obtain

$$\begin{aligned} \epsilon &= \iint_{-1/2T}^{1/2T} K_{xx}(f,v) [e^{j2\pi(f-v)t} - 2 \sum_n e^{j2\pi(ft-vnT)} \phi(t-nT) \\ &\quad + \sum_{n,m} e^{j2\pi(fnT-vmT)} \phi(t-nT) \phi(t-mT)] df dv. \end{aligned}$$

But using the Poisson sum formula [2] we have

$$\sum_n e^{-j2\pi vnT} \phi(t-nT) = \frac{1}{T} \sum_{\ell} \phi(\ell/T-v) e^{-j2\pi(v-\ell/T)t}$$

where

$$\phi(f) = \begin{cases} T, & |f| \leq 1/2T \\ 0, & |f| > 1/2T \end{cases},$$

so that

$$\sum_n e^{-j2\pi vnT} \phi(t-nT) = e^{-j2\pi vt}, \quad |v| \leq 1/2T.$$

Now, substituting this equation into the last expression for ϵ yields

$$\begin{aligned} \epsilon &= \iint_{-1/2T}^{1/2T} K_{xx}(f,v) [e^{j2\pi(f-v)t} - 2e^{j2\pi(f-v)t} + e^{j2\pi(f-v)t}] df dv \\ &= 0 \quad \forall t. \end{aligned}$$

QED

Notice, from Eq. (2-7), that the bandlimiting constraint on $K_{xx}(f,v)$ in the above theorem is satisfied for a process x which is the output of any time-invariant filter whose transfer function is bandlimited to

$(-1/2T, 1/2T)$ regardless of the input process y . Notice also that if x is WSS then the bandlimiting constraint in the above theorem becomes

$$K_{xx}(f, v) = K_{xx}(f) \delta(f-v) = 0$$

if $|f| \geq 1/2T$ or if $|v| \geq 1/2T$ and is satisfied if and only if $K_{xx}(f) = 0$ for $|f| \geq 1/2T$ which is precisely the bandlimiting constraint in the sampling theorem for WSS processes. Hence, the sampling theorem for WSS processes is simply a special case of the general sampling theorem presented here.

(2) Reduction of cyclostationary processes. Any CS process can be reduced to a WSS process via bandlimiting. This fact is stated here as a theorem:

THEOREM(2-3): If x is the output of a filter with transfer function bandlimited to $(-1/2T, 1/2T)$ and the input y is T-CS, then x is WSS.

Proof: From Eq. (2-7) $K_{xx}(f, v)$ is bandlimited to $(-1/2T, 1/2T)$. Thus, from the sampling theorem for deterministic functions [2], we have

$$k_{xx}(t, s) = \sum_{n, m} k_{xx}(nT, mT) \left[\frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \right] \left[\frac{\sin \pi(s-mT)/T}{\pi(s-mT)/T} \right].$$

But, from Theorem (2-1), x is T-CS so that

$$k_{xx}(nT, mT) = k_{xx}[(n-m)T, 0] \quad \forall n, m,$$

and

$$\begin{aligned} k_{xx}(t, s) &= \sum_{m, p} k_{xx}(pT, 0) \left[\frac{\sin \pi((t-mT)-pT)}{\pi((t-mT)-pT)} \right] \left[\frac{\sin \pi(s-mT)/T}{\pi(s-mT)/T} \right] \\ &= \sum_m k_{xx}(t-mT, 0) \left[\frac{\sin \pi(s-mT)/T}{\pi(s-mT)/T} \right] \\ &= k_{xx}(t-s, 0) \quad \forall t, s, \end{aligned}$$

where the last two equalities are due to the sampling theorem for deterministic functions. Now, from Eq. (2-2) it is easily shown that the Fourier transform of $m_x(t)$ is bandlimited to $(-1/2T, 1/2T)$. Thus, from the sampling theorem for deterministic functions, we have

$$\begin{aligned} m_x(t) &= \sum_n m_x(nT) \left[\frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \right] \\ &= m_x(0) \sum_n \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \\ &= m_x(0) \quad \forall t, \end{aligned}$$

where the second equality is due to the fact that x is T-CS (from Theorem (2-1)), and the last equality is due to the sampling theorem for deterministic functions. Hence, both the mean and autocorrelation functions for x are invariant to arbitrary time shifts so that x is WSS.

QED

Notice that this last theorem does not, as might appear, contradict Theorem (2-1) since any WSS process is CS with any choice of period T .

b) Periodically time-varying systems. The subclass, periodically time-varying linear transformations (with period T), frequently--but sometimes inappropriately⁷--termed "periodically time-varying filters" are characterized by the defining relation $g(t, \tau) = g(t+T, \tau+T)$ for all t and τ . That is, the impulse response (kernel) for a linear T -periodic transformation is invariant to time-translations that are integer multiples of the period T .

⁷In the event that the transformation is non-dispersive--it's impulse-response function contains a factor of the form $\delta(f(t)-\tau)$ for some function $f(t)$ --it is not a filter.

For many analyses involving linear time-varying systems (transformations) it is beneficial to separate the time-varying and time-invariant characteristics. The system function defined by Zadeh [35,36] and its inverse Fourier transform do just this. The system function H is defined as follows :

$$H(t,f) \triangleq \int_{-\infty}^{\infty} g(t,t-\tau) e^{-j2\pi f\tau} d\tau, \quad (2-10)$$

and the inverse Fourier transform w.r.t. (with respect to) f of this system function is

$$h(t,\tau) = g(t,t-\tau), \quad (2-11)$$

where τ is frequently referred to as the age variable or memory variable and t may be referred to as the fluctuation variable. Now, the time invariant behavior is characterized by the section function $h(t,\cdot)$ where t is considered to be a parameter, and the time-varying behavior is characterized by the section function $h(\cdot,\tau)$ where τ is a parameter.⁸

For example, if the transformation is time-invariant, then $h(t,\tau) = g(t-(t-\tau)) = g(\tau) = h(0,\tau)$ for every t and there are no time fluctuations-- h depends solely on the age variable. Similarly, if the transformation is T -periodic, then $h(t+T,\tau) = g(t+T,t+T-\tau) = g(t,t-\tau) = h(t,\tau)$ for all t and τ , and the periodicity is reflected only in the fluctuation variable. Treating τ as a parameter, we can expand the T -periodic

⁸A quantity closely related to the system function is the bifrequency function which is simply the Fourier transform w.r.t. t of $H(t,f)$. The bifrequency function finds much application in analyses of time-varying systems, but will not be used here.

section function $h(\cdot, \tau)$ into a Fourier series

$$h(t, \tau) = \sum_{n=-\infty}^{\infty} h_n(\tau) e^{j2\pi n t/T}. \quad (2-12)$$

Thus, from Eq. (2-11), the impulse response g has the representation

$$g(t, \tau) = \sum_n h_n(t-\tau) e^{j2\pi n t/T} \quad (2-13)$$

One interesting result of this representation is the obvious realization of a periodic system as a parallel connection of time-invariant filters $\{h_n\}$ followed by periodic multipliers $\{e^{j2\pi n t/T}\}$ as shown in Figure (2-2). Notice that if the sine-cosine Fourier series were used rather than the exponential Fourier series, then all components--sine and cosine multipliers and time-invariant filters--in the realization would be real not complex. (See Chapter III for other realizations of periodically time-varying linear systems)

The double Fourier transform of the impulse response of a T-periodic system also has an interesting representation which we obtain directly from Eq. (2-13):

$$\begin{aligned} G(f, v) &= \iint_{-\infty}^{\infty} g(t, \tau) e^{-j2\pi(ft-v\tau)} dt d\tau \\ &= \iint_{-\infty}^{\infty} \sum_n h_n(t-\tau) e^{-j2\pi t(f-n/T)} e^{j2\pi \tau v} dt d\tau \\ &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} H_n(f-n/T) e^{-j2\pi \tau(f-n/T)} e^{j2\pi \tau v} d\tau \\ &= \sum_n H_n(f-n/T) \delta(f-v-n/T), \end{aligned} \quad (2-14)$$

where H_n is the single Fourier transform of h_n . We see that $G(f, \nu)$ consists of impulse fences along lines parallel to the $(f = \nu)$ -diagonal, so that the general input-output relation of Eq. (2-5) reduces to:

$$K_{xx}(f, \nu) = \sum_{n=-\infty}^{\infty} \cdot \sum_{m=-\infty}^{\infty} H_n(f-n/T) H_m^*(\nu-m/T) K_{yy}(f-n/T, \nu-m/T), \quad (2-15)$$

Notice that if we consider the degenerate case of a time-invariant filter, then $H_n = 0$ for $n \neq 0$, and $H_0(f) = G(f)$ so that Eq. (2-15) reduces to Eq. (2-9) as it should.

Now, if the input y is WSS then, from Eqs. (2-8), (2-15), the double Fourier transform of the output-process autocorrelation function takes the form

$$\begin{aligned} K_{xx}(f, \nu) &= \sum_{n,m} H_n(f-n/T) H_m^*(\nu-m/T) K_{yy}(f-n/T, \nu-m/T) \\ &\quad \delta(f-\nu-(n-m)/T) \\ &= \sum_p \left[\sum_n H_n(f-n/T) H_{n-p}^*(\nu-p/T) K_{yy}(f-n/T, \nu-p/T) \right] \\ &\quad \delta(f-\nu-p/T). \end{aligned} \quad (2-16)$$

Thus, we see that $K_{xx}(f, \nu)$ also consists of impulse fences on lines parallel to the $(f = \nu)$ -diagonal so that $k_{xx}(t, \tau)$ must exhibit the same periodicity as $g(t, \tau)$. We state this result as a theorem which provides the foundation for the generation of a large class of CS processes:

THEOREM(2-4): If the input y to a T -periodic system is WSS, then the output x is T -CS.

Proof: From Eq. (2-3), we have

$$\begin{aligned}
k_{xx}(t+T, s+T) &= \iint_{-\infty}^{\infty} g(t+T, \tau) g^*(s+T, \gamma) k_{yy}(\tau-\gamma) d\tau d\gamma \\
&= \iint_{-\infty}^{\infty} g(t, \tau-T) g^*(s, \gamma-T) k_{yy}(\tau-T-(\gamma-T)) d\tau d\gamma \\
&= \iint_{-\infty}^{\infty} g(t, \tau') g^*(s, \gamma') k_{yy}(\tau'-\gamma') d\tau' d\gamma' \\
&= k_{xx}(t, s) \quad \forall t, s,
\end{aligned}$$

and, from Eq. (2-2),

$$\begin{aligned}
m_x(t+T) &= \int_{-\infty}^{\infty} g(t+T, \tau) m_y(\tau) d\tau \\
&= \int_{-\infty}^{\infty} g(t, \tau-T) m_y(\tau-T) d\tau \\
&= \int_{-\infty}^{\infty} g(t, \tau') m_y(\tau') d\tau' \\
&= m_x(t) \quad \forall t,
\end{aligned}$$

where g is the impulse response of the T -periodic system. Hence, both the mean and autocorrelation functions for x are T -periodic so that x is T -CS.

QED

We now present four examples of CS processes which can be generated by passing WSS processes through periodic systems. These examples are all signal formats used in communication.

MODEL(1): Amplitude-modulation (AM). Conventional amplitude-modulation is nothing more than multiplication of a signal by a periodic waveform $p(t)$ -- usually a sinusoid. If x is the AM process obtained from the WSS signal y as follows: $x(t) = p(t)y(t)$, then we see that x is obtained from a periodic linear transformation on y where the kernel (impulse response) is $g(t, \tau) = p(t)\delta(t-\tau)$. Thus, from Theorem (2-4), x is CS. An interesting

special case of AM is that where y has a zero-mean value and $p(t)$ is a sinusoid with period T . Then

$$\begin{aligned} k_{xx}(t,s) &= \cos(2\pi t/T) \cos(2\pi s/T) k_{yy}(t-s) \\ &= \frac{1}{2} [\cos(2\pi(t+s)/T) + \cos(2\pi(t-s)/T)] k_{yy}(t-s), \end{aligned}$$

so that x is CS with period $T/2$ even though $p(t)$ is T -periodic, not $T/2$ -periodic.

MODEL(2): Pulse-amplitude-modulation (PAM). As discussed in Chapter I, a synchronous PAM signal x is simply a periodic train of equally spaced (in time) pulses with random amplitudes:

$$x(t) = \sum_{n=-\infty}^{\infty} a_n q_0(t-nT).$$

The random sequence $\{a_n\}$ may be a random data sequence with a finite alphabet (finite number of admissible values for realizations of each a_n) or $\{a_n\}$ may be the sample values $\{y(nT)\}$ of some continuous-time process y . In this latter case, x can be generated by passing y through a periodic system composed of a periodic impulse sampler followed by a time-invariant filter with impulse response $q_0(t)$ as shown in Figure (2-3). The impulse response for this composite T -periodic system is

$$g(t, \tau) = q_0(t-\tau) \sum_n \delta(\tau-nT).$$

If y is WSS then, from Theorem (2-4) x is T -CS.

In Sec. 7 of this chapter we discuss a generalized model of PAM which includes a random disturbance termed jitter. In this generalized model, the pulse stream is no longer synchronous, but x is still CS. Also, if the pulses comprising a single PAM signal are sufficiently

separated in time, then several of these signals can be interleaved in time to form a composite time-division-multiplex PAM signal which will be CS if the random amplitude sequences are jointly WSS. Note that processes which are individually WSS and are statistically independent are also jointly WSS.

MODEL(3): Time-division-multiplex (TDM_1). The TDM_1 process, as discussed in Chapter I, is formed by interleaving finite-length records from a multiplicity, say M , of signal sources. This can be done by passing each component process through a periodic gate, and summing the gated processes as shown in Figure (2-4a). A typical periodic gate-function is shown in Figure (2-4b). Realizations of the composite process take the form

$$x(t) = \sum_{n=-\infty}^{\infty} y_n(t) q_0(t-nT_0)$$

where $T_0 = T/M$ and $y_n = y_{n+M}$ for all n . As pointed out in Chapter I, x is CS with period T if the $\{y_n\}$ are jointly WSS. This result can be explained in terms of a vector formulation of Theorem (2-4) where the WSS y is replaced by an M -vector of jointly WSS processes, and the T -periodic system is replaced with a $1 \times M$ matrix of T -periodic systems.

Alternatively, the cyclostationarity of x can be attributed to the fact that a sum of jointly T -CS processes is a T -CS process.

Notice that the gating operation in the multiplexor discards portions of each component process, so that conceivably there could be a loss of information in this multiplexing scheme. In the following subsection on time-scale transformations, we introduce a "lossless" time-division multiplexing scheme, TDM_2 .

MODEL(4): Frequency-division-multiplex (FDM). A multiplexing scheme which is frequently encountered in radio communication systems is that of frequency-division-multiplex (FDM). The FDM signal is formed by interleaving in frequency a multiplicity, say M , of "frequency-translated" signals, and is to some extent the frequency-time dual of TDM. Specifically, each component process y_n is bandlimited to $(-1/2T, 1/2T)$ and then amplitude-modulated with a sinusoid (translated in frequency) into disjoint frequency bands, as shown in Figure (2-5b), and the resultant translates are summed to form the composite FDM signal with realizations of the form:

$$x(t) = \sum_{n=1}^M x_n(t) \cos(\omega_0 t + 2\pi n t / T),$$

where the $\{x_n\}$ are the bandlimited versions of the $\{y_n\}$. The complete multiplexor is shown in Figure (2-5a).

Now, if ω_0 is an integer multiple of $2\pi/T$, and the $\{x_n\}$ are jointly WSS, then x is the sum of M jointly T-CS processes and is therefore T-CS. Again, as in the case of TDM, this result can be interpreted in terms of a vector version of Theorem (2-4).

Notice that in both TDM₁ and FDM, the composite signal is the sum of CS AM processes; that is, the sum of deterministic periodic signals amplitude-modulated by random WSS signals.

c) Time-scale transformations. The realization $y(t')$ of a process which has undergone a time-scale transformation $t'(t)$ is a function of a function $y[t'(t)]$ and may therefore be referred to as a composition. We will consider WSS input processes y which undergo time-scale transformations

consisting of a scaled identity function and a periodic function $p(t)$, so that the time parameter t is replaced by $t' = at + p(t)$ to yield an output composition-process x with realizations of the form

$$x(t) = y(at+p(t)). \quad (2-17)$$

Time-scale transformations of this form can be characterized by their impulse response functions which take the form

$$g(t, \tau) = \delta(\tau - at - p(t)) \quad (2-18)$$

Notice that

$$\begin{aligned} g(t+T, \tau+T) &= g(t, \tau + (1-a)T) \\ &\neq g(t, \tau) \quad \text{if } a \neq 1, \end{aligned}$$

so that these "periodic" time-scale transformations are not periodic systems unless $a = 1$.

The following theorem serves as a basis for the generation of a large class of CS processes:⁹

THEOREM(2-5): If x is a process with realizations obtained from the realizations of a WSS process y by the time-scale transformation: $x(t) = y(at+p(t))$, where a is a deterministic constant and $p(t)$ is a deterministic T -periodic function, then x is T -CS and has constant mean and variance.

Proof:

$$\begin{aligned} m_x(t+T) &= E\{y(a(t+T)+p(t+T))\} \\ &= m_y, \text{ a constant} \quad \forall t \\ &= m_x(t) \quad \forall t, \end{aligned}$$

⁹This class has also been considered by Hurd [5].

and

$$\begin{aligned}
 k_{xx}(t+T, s+T) &= E\{y(a(t+T)+p(t+T))y(a(s+T)+p(s+T))\} \\
 &= E\{y(a(t+T)+p(t))y(a(s+T)+p(s))\} \\
 &= k_{yy}(a(t+T)+p(t)-a(s+T)-p(s)) \\
 &= k_{yy}(at+p(t)-as-p(s)) \\
 &= k_{xx}(t, s),
 \end{aligned}$$

so that x is T-CS with constant mean, and variance

$$\sigma^2 = k_{xx}(t, t) - m_x^2(t) = k_{yy}(0) - m_y^2$$

which is also constant.

QED

It is worth emphasizing that all members of the subclass of CS processes which are generated from WSS processes via periodic time-scale transformations have constant mean and variance regardless of the severity of periodic fluctuations in the transformations.

We proceed with three examples:

MODEL(5): Time-division-multiplex (TDM_2). The time-division-multiplexing scheme (TDM_1) introduced in Chapter I and discussed earlier in this section can result in a loss of information in the sense that only a fraction $(1/M)$ of every T -length record of the component processes is contained in the composite TDM_1 process--the remaining fraction $((M-1)/M)$ is discarded by the gates in the multiplexor (see Fig. (2-4)). This loss of information can be circumvented if every T -length record is first compressed in time to a T/M -length record. The resultant periodically compressed component

processes can then be multiplexed together without loss of information.

The periodic time-scale compression can be effected by applying a time-scale transformation to the original component processes:

$x_n(t) = y_n(t+p(t))$. An appropriate function $p(t)$ is the periodic triangular waveform shown in Figure (2-6a). With this choice, the transformed time variable $t' = t + p(t)$ remains constant (time does not advance) over the first $(M-1)/M$ of each adjacent T -length interval, and then increases linearly with slope M (time advances at a rate accelerated by the factor M) over the final $1/M$ of each adjacent T -length interval (see Figure (2-6b)).

The resultant effect of this periodic time-scale compression is shown in Figure (2-7) with a typical realization of the M th process y_M and its compressed counterpart x_M .

In order that the M processes, to be time-division-multiplexed, can be properly interleaved, the component processes must be compressed into interleaving time-slots as expressed here:

$$x_n(t) = y_n(t+p(t+(M-n)T/M)).$$

Now, the compressed component processes can be applied to the multiplexor of Figure (2-4) to yield the composite TDM_2 signal with no loss of information, and with realizations of the form:

$$x(t) = \sum_n y_n(t+p(t+(M-n)T_0))q_0(t-(n-1)T_0)$$

where $T_0 = T/M$.

If the original component processes $\{y_n\}$ are jointly WSS, then each compressed component process x_n will be T-CS (from Theorem (2-5)) and, in fact, the M components will be jointly T-CS. An extended version of Theorem (2-4) (a vector of jointly T-CS inputs into a matrix of T-periodic systems yields a T-CS output) guarantees that the composite TDM₂ signal x is T-CS.

Notice that any time-scale transformation (including the time-scale compression in the above example) can be realized, in principle at least, with a recording machine which either records or plays back (not both) at a variable speed determined by the time-scale transformation.

MODEL(6): Doppler shifted processes. A realization of any random waveform--be it acoustic, electromagnetic, mechanical, etc.--which is transmitted, radiated, or reflected from a body whose spatial frame of reference varies w.r.t. that of an observer of the random waveform will undergo a time-scale transformation which is frequently referred to as a Doppler effect or Doppler shift. If the spatial variation is an oscillation or is, in some sense, periodic so that the resultant time-scale transformation exhibits the periodicity described in the hypothesis of Theorem (2-5), and if the original process is WSS, then the resultant Doppler-shifted process is CS (by Theorem (2-5)).

This Doppler effect is encountered frequently in studies in nuclear, atomic, and astronomical physics, and in radio-science, radio-astronomy, and other fields of science.

i) Processes derived from the Poisson counting process with periodic rate parameter: In contrast to the synchronous pulse-train processes such as PAM, one often encounters processes where only the average rate of pulse occurrences is known. In such situations, where pulse occurrences are

totally asynchronous (random), the Poisson counting process is frequently used in constructing a model [2,3,22]. In particular, models for random facsimile signals [2], random telegraph signals [2,3], and shot noise [2,3,22] have been constructed on the basis of the Poisson counting process.

If the rate parameter μ in the counting process is constant, then the above models are WSS. However, if the rate parameter varies periodically with time, then the models become CS. Such CS models have been proposed in studies of meteorology [18] and noise in electronic devices [22].

We will show here that the mean and autocorrelation functions for the CS random facsimile signal derived from the inhomogeneous Poisson counting process with rate parameter $\mu(t) = \mu_0 + \mu_1 p(t)$ are identical to those corresponding to the process which is obtained from the time-scale transformation: $t' = t + \frac{\mu_1}{\mu_0} \int p(t)$ on the WSS facsimile signal derived from the homogeneous Poisson counting process with constant rate parameter μ_0 . (A similar, but less general result, for the random telegraph signal has been stated by Hurd [5].)

MODEL(7): Random facsimile signal. A random facsimile signal is obtained by optically scanning a black and white picture according to, for example, the line scanning scheme discussed in the following subsection. For a random picture, the scanner output x will be a two-valued random process, say $x(t) = 1$ for black portions and $x(t) = 0$ for white portions. The times at which transitions from black-to-white and white-to-black occur will be an ordered sequence of random variables which we will model in terms of the inhomogeneous Poisson counting process with

rate parameter μ . Using precisely the same assumptions and method of derivation as Franks [2], but generalized from the case of a constant rate parameter μ_0 to that of the fluctuating $\mu(t)$, we obtain the following set of coupled differential equations:¹⁰

$$\begin{aligned} \frac{\partial P_n(t,s)}{\partial s} &= -\mu(s)[P_n(t,s) - P_{n-1}(t,s)], \quad s \geq t, \quad n \geq 0 \\ P_n(t,t) &= \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \\ P_n(t,s) &= 0, \quad n < 0, \end{aligned} \quad (2-19)$$

where $P_n(t,s)$ is defined to be the probability that exactly n transitions occur in the interval (t,s) . x is constant over each interval defined by consecutive transition times, and assumes the values 1 or 0 with probabilities p or $1-p$, and the values of x in different intervals are assumed to be statistically independent.

Now, from this model, we find that:

$$\begin{aligned} m_x(t) &= E\{x(t)\} = \Pr(x(t)=1) = p \quad \forall t, \\ k_{xx}(t,s) &= E\{x(t)x(s)\} = \Pr(x(t)=1 \text{ and } x(s)=1) \\ &= p \cdot \Pr(t \text{ and } s \text{ are in the same interval}) \\ &\quad + p^2 \cdot \Pr(t \text{ and } s \text{ are in different intervals}) \\ &= p \cdot P_0(t,s) + p^2(1 - P_0(t,s)) \quad \forall s \geq t, \end{aligned}$$

where $\Pr(\cdot)$ indicates "the probability of the event in the parenthesis".

But, from Eq. (2-19), we easily obtain the solution for P_0 :

$$P_0(t,s) = \exp\left[-\int_t^s \mu(\sigma) d\sigma\right], \quad s \geq t,$$

¹⁰See also Parzen [22], and Papoulis [3].

so that

$$k_{xx}(t,s) = (m_x - m_x^2) \exp\left[-\int_t^s \mu(\sigma) d\sigma\right] + m_x^2 \quad \forall s \geq t,$$

and since $k_{xx}(t,s)$ is symmetrical in t and s then

$$k_{xx}(t,s) = (m_x - m_x^2) \exp\left[-\left|\int_t^s \mu(\sigma) d\sigma\right|\right] + m_x^2 \quad \forall t,s.$$

If we now consider a process y with constant rate parameter μ_0 , then

$$k_{yy}(t,s) = (m_x - m_x^2) \exp[-\mu_0 |t-s|] + m_x^2, \quad \forall t,s$$

which is the result derived by Franks [2]. The interesting point here is that if we subject y to the time-scale transformation: $t' = t + \frac{\mu_1}{\mu_0} \int p(t)$, then we obtain a process x' with the same mean and autocorrelation functions as the above process x with fluctuating rate parameter $\mu(t) = \mu_0 + \mu_1 p(t)$.

We conjecture that various WSS models for piecewise constant processes derived from the Poisson counting process with constant rate parameter μ_0 can be extended to CS models with the time-scale transformation: $t' = t + \frac{\mu_1}{\mu_0} \int p(t)$, where the resultant mean and autocorrelation functions will be identical to those corresponding to the model obtained directly in terms of the inhomogeneous Poisson counting process with periodic rate parameter: $\mu(t) = \mu_0 + \mu_1 p(t)$.

MODEL(8): Asynchronous PAM. A specific model which we will be concerned with in Chapter IV can be derived, via a time-scale transformation, from the general WSS asynchronous random pulse-sequence

$$y(t) = \sum_{n=-\infty}^{\infty} a_n q_0(t-t_n) \quad (2-20)$$

where $\{a_n\}$ is an arbitrary WSS sequence independent of the occurrence times $\{t_n\}$ which form an ordered sequence distributed according to the Poisson counting process with constant rate parameter μ_0 , and $q_0(t)$ is an arbitrary pulse shape. This process y has been used as a model for shot noise and, as shown by Franks [2], has mean and autocorrelation functions given by

$$m_y = \mu_0 m_a \int_{-\infty}^{\infty} q_0(t) dt$$

$$k_{yy}(\tau) = m_y^2 + \mu_0 \sigma_a^2 \int_{-\infty}^{\infty} q_0(t+\tau) q_0(t) dt,$$

where σ_a^2 is the variance of $\{a_n\}$ and m_a is the mean. Now, we create a new process x by subjecting y to the time-scale transformation:

$t' = t + p(t)$, where $p(t)$ is any T -periodic function such that $t'(t)$ is a nondecreasing function (time never regresses), and obtain

$$m_x(t) = m_y$$

$$k_{xx}(t, s) = k_{yy}(t + p(t) - s - p(s))$$

as the mean and autocorrelation functions for the T -CS asynchronous random pulse-sequence x whose average rate of pulse-occurrences (and pulse-shapes) varies periodically with time.

3. Multi-dimensional Linear Transformations (Scanning)

In this section we briefly discuss the transformation of multi-dimensional quasi-stationary processes into one-dimensional cyclostationary processes, and the application of this concept to the generation of CS processes via scanning.

A three-dimensional process y is a random function of three independent variables: $y(t,u,v)$, $t \in R$, $u \in D_1$, $v \in D_2$. In the applications discussed here, u and v represent two spatial coordinates and t the usual time coordinate. The domains D_1 and D_2 will be finite intervals in R the set of real numbers. We will define y to be quasi-wide-sense stationary if and only if each of the one-dimensional processes $y(\cdot, u, v)$ is WSS for each value of $u \in D_1$ and $v \in D_2$.¹¹

We are interested in linear transformations which map y into a one-dimensional CS process x (a random function of the single variable t). As a generalization of the one-dimensional linear transformations discussed in Section 2 of this chapter, we characterize a three-dimensional linear transformation in terms of a three-dimensional impulse response.¹² If y is the input to a transformation G , and x is the output, then the realizations of x are related to the corresponding realizations of y as follows:

$$x(t) = \int_{D_1} \int_{D_2} \int_R g(t, \tau, u, v) y(\tau, u, v) d\tau dv du, \quad (2-21)$$

¹¹This notation does not parallel our convention for one-dimensional processes, but is the least ambiguous.

¹²The terms "three-dimensional linear transformation" and "three-dimensional impulse response" are not quite appropriate since the output (image) process is one-dimensional.

where $g(t, \tau, u, v)$ is the three-dimensional impulse response.

We say that G is a quasi-periodic transformation with period T if and only if

$$g(t+T, \tau+T, u, v) = g(t, \tau, u, v) \quad \forall t, \tau \in \mathbb{R}, u \in D_1, v \in D_2. \quad (2-22)$$

With these definitions we can generalize Theorem (2-4) as follows:

THEOREM(2-6): If the input y to a three-dimensional quasi-periodic transformation with period T is a three-dimensional quasi-WSS process, then the one-dimensional output x is a T -CS process.

The proof directly parallels that for Theorem (2-4).

We now give two applications of Theorem (2-6) to the generation of CS processes via scanning:

MODEL(9): Line-scanning a rectangular field (video). The process of scanning a finite rectangular field--performed by a video camera, for example--can be modeled with a three-dimensional linear transformation. Then the scanner is fully characterized by the impulse response $g(t, \tau, u, v)$ which is sometimes referred to as a window function. With reference to Figure (2-8) and Eq. (2-21), $D_1 = (0, W_1)$ and $D_2 = (0, W_2)$ are the domains for u and v , and $y(t_0, u_0, v_0)$ is a realization of the three-dimensional process y evaluated at time t_0 , horizontal distance u_0 , and vertical distance v_0 . The video signal is the scanner output process x and is obtained from y as in Eq. (2-21).

The conventional video camera employs a constant velocity line-scanner whose impulse response can be modeled as (with some abuse of notation):

$$g(t, \tau, u, v) = g(u_1(t) - u, v_1(t) - v) \delta(t - \tau)$$

where the functions u_1 and v_1 are as shown in Figure (2-9), and each are LT -periodic where T is the time required to scan a single line and L is the number of lines per frame. Therefore, the scanner G is a quasi-periodic transformation with period LT , so that if y is quasi-WSS then the video process x is CS (period LT), by Theorem (2-6).

If the above window function $g(\cdot, \cdot)$ is ideal--a two-dimensional impulse function--then we have

$$x(t) = y(t, u_1(t), v_1(t)).$$

Using this ideal model for the scanner along with the assumption of zero retrace time (see Fig. (2-9)), and the model for y proposed by Franks [10], results in an autocorrelation function for x which is composed of three periodic factors. The first factor characterizes the correlation within any given line, the second characterizes line-to-line correlation, and the third characterizes frame-to-frame correlation as follows:

$$k_{xx}(t, s) = k_1(t-s)k_2(t, s)k_3(t, s) \quad (2-23)$$

where, for

correlation within a line:

$$\begin{aligned} k_1(t-s) &= \rho_1^{(t-s) \bmod T} \\ &= \sum_{n=-\infty}^{\infty} \rho_1^{|t-s-nT|} q_0(t-s-nT) \end{aligned} \quad (2-24)$$

and is T -periodic in $(t-s)$ (q_0 is a rectangular gate function of width T),

line-to-line correlation:

$$\begin{aligned}
 k_2(t,s) &= \rho_2 \frac{1}{T} |(t-t \bmod T) \bmod LT - (s-s \bmod T) \bmod LT| \\
 &= \sum_{n,m=-\infty}^{\infty} \rho_2^{(n-m) \bmod L} q_0(t-nT) q_0(s-mT)
 \end{aligned} \tag{2-25}$$

and is jointly T-periodic in t and s,

frame-to-frame correlation:

$$\begin{aligned}
 k_3(t,s) &= \rho_3 \frac{1}{T} |(t-t \bmod LT) - (s-s \bmod LT)| \\
 &= \sum_{n,m} \rho_3^{|n-m|} q_1(t-nLT) q_1(s-mLT)
 \end{aligned} \tag{2-26}$$

and is jointly LT-periodic in t and s (q_1 is a rectangular gate function of width LT).

Finally, if the contribution to $k_{xx}(t,s)$ due to frame-to-frame correlation is ignored because of its relatively low-frequency nature (this is done in practice for some analyses of video systems) then the period of cyclostationarity reduces from LT to T as in the model introduced in Chapter I.

MODEL(10): Circular scanning (radar). The process of continuously scanning a one-dimensional circular field--performed by a radar antenna, for example--can be modeled with a two-dimensional linear transformation with impulse response $g(t,\tau,\theta)$. With reference to Figure (2-10) and the two dimensional analog of Eq. (2-21), $D_1 = (0,2\pi)$ is the domain for θ , and $y(t_0, \theta_0)$ is a realization of the two-dimensional random process y evaluated at time t_0 and angular displacement θ_0 . The received radar signal is the scanner output process x and is obtained from y as in

Eq. (2-21):

$$x(t) = \int_R \int_{D_1} g(t, \tau, \theta) y(\tau, \theta) d\theta d\tau \quad (2-27)$$

If a constant velocity scanner is used, then the impulse response can be modeled as (with some abuse of notation):

$$g(t, \tau, \theta) = g(\theta_1(t) - \theta) \delta(t - \tau)$$

where the function θ_1 is as shown in Figure (2-11), and is T -periodic, where T is the time required for one complete revolution of the scanning antenna. Therefore G is a quasi-periodic transformation with period T , so that if y is quasi-WSS then the radar process x is T -CS by the two-dimensional version of Theorem (2-6).

As an example of a "non-ideal" window function g , consider the rectangular window shown in Figure (2-12). For this case, the signal x is given by

$$\begin{aligned} x(t) &= \int_0^{2\pi} \int_{-\infty}^{\infty} g(\theta_1(t) - \theta) \delta(t - \tau) y(\tau, \theta) d\tau d\theta \\ &= \int_0^{2\pi} g(\theta_1(t) - \theta) y(t, \theta) d\theta \\ &= \frac{5}{\pi} \int_{\theta_1(t) - \frac{\pi}{10}}^{\theta_1(t) + \frac{\pi}{10}} y(t, \theta) d\theta \\ &= \frac{5}{\pi} \int_{\frac{2\pi t}{T} - \frac{\pi}{10}}^{\frac{2\pi t}{T} + \frac{\pi}{10}} y(t, \theta) d\theta \end{aligned}$$

and is the average value of $y(t, \theta)$ over the range $(-\pi/10, \pi/10)$ of θ .

In Section 6, we generalize the discussion here to random multi-dimensional linear transformations, and show that random scanners with random window shapes and velocities can also generate cyclostationary signals.

4. Nonlinear Transformations

Many nonlinear transformations which arise in engineering problems can be characterized with Wiener's Volterra series representation [44-47] in which the output $x(t)$ of an N^{th} order Volterra system driven by the input $y(t)$ is expressed as

$$x(t) = \sum_{n=1}^N \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_n(t-\tau_1, t-\tau_2, \dots, t-\tau_n) y(\tau_1) \cdots y(\tau_n) d\tau_1 \cdots d\tau_n,$$

where the kernels $\{g_n\}$ can be thought of as generalized impulse response functions. This series is sometimes referred to as a power series with memory.

Note that a first order Volterra system ($N = 1$) is simply a linear system and the above representation is, for this case, the super-position integral (Eq.(2-1)) used in Section 2 for representing linear transformations. Hence g_1 is for this case, the usual impulse-response function.

The nonlinear transformations of interest in this thesis can, for the most part, be characterized with a generalized Volterra series in which the kernels are invariant under joint shifts, in their arguments, which are multiples of a basic period T :

$$g_n(t, \tau_1; t, \tau_2; \dots; t, \tau_n) = g_n(t+T, \tau_1+T; \dots; t+T, \tau_n+T), \quad \forall n$$

for all t, τ_1, \dots, τ_n ,

whereas, the kernels in the standard Volterra series are invariant under arbitrary joint shifts.

In addition to allowing periodically time-varying kernels, we will also include a T -periodic zero th order term so that our generalized periodic volterra series becomes:

$$x(t) = g_0(t) + \sum_{n=1}^N \int \cdots \int g_n(t, \tau_1; \cdots; \tau_n) y(\tau_1) \cdots y(\tau_n) d\tau_1 \cdots d\tau_n \quad (2-28)$$

We will refer to this more general class of Volterra systems as periodic Volterra systems. For this class, we have the following generalization of Theorem (2-4) on the generation of CS processes:

THEOREM(2-7):¹³ The output process x of an N^{th} order T -periodic Volterra system, which is driven by a stationary - of order $2N$ ($S(2N)$) input process y , is T -CS.

Proof:

i) From Eq. (2-28), the mean of the output x is

$$\begin{aligned} m_x(t) &= g_0(t) + \sum_{n=1}^N \int \cdots \int g_n(t, \tau_1; \cdots; \tau_n) E\{y(\tau_1) \cdots y(\tau_n)\} d\tau_1 \cdots d\tau_n \\ &= g_0(t+T) + \sum_{n=1}^N \int \cdots \int g_n(t+T, \tau_1+T; \cdots; \tau_n+T) \cdot \\ &\quad E\{y(\tau_1) \cdots y(\tau_n)\} d\tau_1 \cdots d\tau_n \\ &= g_0(t+T) + \sum_{n=1}^N \int \cdots \int g_n(t+T, \tau'_1; \cdots; \tau'_n) \cdot \\ &\quad E\{y(\tau'_1-T) \cdots y(\tau'_n-T)\} d\tau'_1 \cdots d\tau'_n \\ &= g_0(t+T) + \sum_{n=1}^N \int \cdots \int g_n(t+T, \tau'_1; \cdots; \tau'_n) E\{y(\tau'_1) \cdots y(\tau'_n)\} d\tau'_1 \cdots d\tau'_n \\ &= m_x(t+T) \quad \forall t. \end{aligned}$$

where the next-to-last equality is a result of the $S(2N)$ property of y .

¹³A process is $S(N)$ iff its 1st through N^{th} order joint PDF's are invariant under arbitrary joint shifts in their indexing arguments [3].

ii) Again, using Eq. (2-28), we have

$$\begin{aligned}
 k_{xx}(t,s) = & g_0(t)g_0(s) + g_0(t)m_x(s) + g_0(s)m_x(t) \\
 & + \sum_{n=1}^N \sum_{m=1}^N \int \cdots \int g_n(t, \tau_1; \cdots; t, \tau_n) g_m(s, \gamma_1; \cdots; s, \gamma_m) \cdot \\
 & E\{y(\tau_1) \cdots y(\tau_n) y(\gamma_1) \cdots y(\gamma_m)\} d\tau_1 \cdots d\tau_n d\gamma_1 \cdots d\gamma_m
 \end{aligned}$$

Using the periodicity of the $\{g_n\}$ and of m_x , and the stationarity of y , as in part i), yields the result:

$$k_{xx}(t,s) = k_{xx}(t+T, s+T) \quad \forall t, s.$$

Hence, from i) and ii), x is T-CS.

QED

MODEL(11): Parametric amplifiers. The parametric amplifier [48] is an example of a nonlinear system which can be accurately modeled as a periodically time-varying nonlinear transformation with memory, and might therefore be modeled as a periodic Volterra system. Hence, the output of a parametric amplifier driven by a strict-sense-stationary (SSS) process¹⁴ might well be expected to be CS, and has, in fact, been so modeled in Reference [21].

In the next two subsections, we consider special periodic Volterra systems which generate CS process from input processes which are only S(2).

a) Zero-memory nonlinearities. If the kernel functions in the periodic Volterra series representation of a system take the form

$$g_n(t, \tau_1; t, \tau_2; \cdots; t, \tau_n) = g_n(t) \delta(t - \tau_1) \cdots \delta(t - \tau_n)$$

¹⁴A process is SSS iff it is S(N) for all natural numbers N [3].

where

$$g_n(t+T) = g_n(t),$$

then we have the reduced input-output relation

$$x(t) = \sum_{n=1}^N g_n(t) (y(t))^n,$$

and the system is a zero-memory periodic nonlinearity. The input-output relation for these systems can be compactly expressed as

$$x(t) = G(y(t), t)$$

where $G(\cdot, t)$ is the nonlinear function with power series coefficients $\{g_n(t)\}$, and

$$G(\cdot, t+T) = G(\cdot, t) \quad \forall t.$$

For these systems, the output x at any time, say t_0 , depends on the input y only at time t_0 , and on the time t_0 itself.

For this class of zero-memory nonlinear systems, we have the following theorem on the generation of CS processes:

THEOREM(2-8): If the input y to a zero-memory time-varying nonlinearity, which satisfies the periodicity condition: $G(\cdot, t) = G(\cdot, t+T)$ for all t , is $S(2)$ then the output x is T -CS.

Proof:

$$\begin{aligned} m_x(t) &= E\{G(y(t), t)\} = \int_{-\infty}^{\infty} G(\sigma, t) p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} G(\sigma, t+T) p(\sigma) d\sigma \\ &= m_x(t+T) \quad \forall t, \end{aligned}$$

where $p(\cdot)$ is the PDF for $y(\cdot)$ for all t ; and

$$\begin{aligned}
k_{xx}(t,s) &= E\{G(y(t),t)G(y(s),s)\} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\sigma,t)G(\gamma,s)p_{t-s}(\sigma,\gamma)d\sigma d\gamma \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\sigma,t+T)G(\gamma,s+T)p_{(t+T)-(s+T)}(\sigma,\gamma)d\sigma d\gamma \\
&= k_{xx}(t+T,s+T) \quad \forall t,s,
\end{aligned}$$

where $p_{t-s}(\cdot,\cdot)$ is the joint PDF for $y(t)$ and $y(s)$. Hence x is T-CS.

QED

As an example of a transformation which satisfies the hypothesis of Theorem (2-8), consider the form:

$$G(y(t),t) = G(y(t) + q(t))$$

where $q(t)$ is any T-periodic function. Here, the signal y is subjected to an additive periodic component $q(t)$ and then applied to a nonlinear device such as a diode, square-law device, or some type of threshold device. This type of transformation is frequently incurred in optimum detectors for periodic signals in stationary noise. Note that if G were a square-law device, then the generalized Volterra kernels would be given explicitly by

$$g_n = 0 \quad n \geq 3$$

$$g_2(t, \tau_1; t, \tau_2) = \delta(t - \tau_1)\delta(t - \tau_2)$$

$$g_1(t, \tau_1) = 2q(t)\delta(t - \tau_1)$$

$$g_0(t) = q^2(t).$$

b) Nonlinear modulation of synchronous pulse-trains. If the kernel functions in the periodic Volterra series representation take the form

$$g_n(t, \tau_1; \dots; t, \tau_n) = \sum_{i=-\infty}^{\infty} q_n(t-iT) \delta(iT-\tau_1) \dots \delta(iT-\tau_n),$$

then we have the reduced input-output relation

$$x(t) = \sum_{i=-\infty}^{\infty} \left[\sum_{n=1}^N g_n(t-iT) (y(iT))^n \right]$$

which can be expressed in the more compact form

$$x(t) = \sum_{i=-\infty}^{\infty} q(t-iT, y(iT)),$$

where $q(t, \cdot)$ is the nonlinear function with power series coefficients $\{g_n(t)\}$. This last expression represents a signal format which has received much attention as a means for transmitting random sequences of numbers $\{a_i\} = \{y(iT)\}$, and is referred to as a modulated synchronous pulse-train.

With this signal format, each pulse $q(t-iT, \cdot)$ carries the realization of one random number, a_i . If the modulation is linear, then the signal is PAM, as discussed in Section 2b, and the random numbers are carried as pulse-amplitudes. However, if the modulation is nonlinear, then there is a multitude of different schemes--each with its own theoretical and practical advantages and disadvantages. Following are six specific examples:

MODEL(12): Frequency-shift-keying (FSK). A sinusoidal signal with gating envelope $w(t)$ is frequency-shifted by the parameters $\{a_n\}$ such that

$$\begin{aligned} q(t-nT, a_n) &= w(t-nT) \cos(2\pi(f+a_n)(t-nT)) \\ &= w(t-nT) \cos(2\pi(f+a_n)t) \end{aligned}$$

where the second equality is valid if $(f+a_n)T$ is an integer for every realization of a_n for every n ,

MODEL(13): Phase-shift-keying (PSK). A sinusoidal signal with gating envelope $w(t)$ is phase-shifted by the parameters $\{a_n\}$ such that

$$\begin{aligned} q(t-nT, a_n) &= w(t-nT) \cos(2\pi f(t-nT) + a_n) \\ &= w(t-nT) \cos(2\pi f t + a_n) \end{aligned}$$

where the second equality is valid if fT is an integer.

MODEL(14): Pulse-width-modulation (PWM). The width (duration) of a basic pulse shape $q_0(t)$ is determined by the parameters $\{a_n\}$ such that

$$q(t-nT, a_n) = q_0((t-nT)/a_n)$$

where $\{a_n\}$ take on only positive values.

MODEL(15): Pulse-position-modulation (PPM). The epoch (time-of-occurrence) of a basic pulse $q_0(t)$ is determined by the parameters $\{a_n\}$ such that

$$q(t-nT, a_n) = q_0(t-nT-a_n)$$

where $\{a_n\}$ take on only values which are non-negative and less than T .

MODEL(16): Digital pulse-frequency-modulation (DPFM). The frequency-of-occurrence--number--of basic pulses $q_0(t)$ in each time-slot is determined by the parameters $\{a_n\}$ such that

$$q(t-nT, a_n) = \sum_{i=1}^{a_n} q_0((t-nT)a_n - (i-1)T),$$

where $\{a_n\}$ take on only non-negative integer values. Note that if $q_0(t)$ is an integral number of cycles of a sinusoid with duration T , then this DPFM signal is an FSK signal with rectangular envelope $w(t)$.

MODEL(17): General M -ary signaling. The parameters $\{a_n\}$ serve to index a set of M basic pulse-shapes, which need not be "functionally" related as in the above cases, such that

$$q(t-nT, a_n) = q_{a_n}(t-nT)$$

where, for example, $\{q_i(t); i = 1, 2, 3, \dots, M\}$ might be a set of mutually orthogonal pulses duration limited to the interval $(0, T)$.

Paralleling Theorem (2-8), we have the following theorem which establishes a sufficient condition for the general synchronous pulse-train signal to be CS:

THEOREM(2-9): If x is a synchronous pulse-train with period T and the modulating sequence $\{a_n\}$ is stationary-of-order-two, then x is T -CS.

Proof:

$$\begin{aligned} x(t) &= \sum_n q(t-nT, a_n) \text{ so that} \\ m_x(t) &= \int_{-\infty}^{\infty} \sum_n q(t-nT, \sigma) p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} \sum_n q(t+T-(n+1)T, \sigma) p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} \sum_{n'} q(t+T-n'T, \sigma) p(\sigma) d\sigma \\ &= m_x(t+T) \quad \forall t, \end{aligned}$$

where $p(\cdot)$ is the PDF for a_n for every n , and

$$\begin{aligned}
k_{xx}(t,s) &= \iint_{-\infty}^{\infty} \sum_{n,m} q(t-nT,\sigma)q(s-mT,\gamma)p_{n-m}(\sigma,\gamma)d\sigma d\gamma \\
&= \iint_{-\infty}^{\infty} \sum_{n,m} q(t+T-(n+1)T,\sigma)q(s+T-(m+1)T,\gamma)p_{(n+1)-(m+1)}(\sigma,\gamma)d\sigma d\gamma \\
&= \iint_{-\infty}^{\infty} \sum_{n',m'} q(t+T-n'T,\sigma)q(s+T-m'T,\gamma)p_{n'-m'}(\sigma,\gamma)d\sigma d\gamma \\
&= k_{xx}(t+T,s+T) \quad \forall t,s,
\end{aligned}$$

where $p_{t-s}(\cdot,\cdot)$ is the joint PDF for $x(t)$ and $x(s)$. Hence x is T-CS.

QED

In Section 7 we present a generalized model for pulse-train signals which incorporates random jitter wherein the pulse train is no longer synchronous but the signal is still CS.

5. Random Linear Transformations

The developments in this section on random linear transformations closely parallel those in Section 2 on deterministic linear transformations. Due to the length of this section, we begin by giving a brief outline of the theorems proved and the cyclostationary process models defined:

a) Wide-sense-stationary systems

Theorem(2-10): Preservation of wide-sense stationarity by random transformations.

Model(18): Random periodic signals

Theorem(2-11): Conversion of deterministic periodic systems into wide-sense-stationary systems via phase-randomization.

Theorem(2-12): Preservation of cyclostationarity by random transformations.

b) Cyclostationary systems

Theorems(2-13): Generation of cyclostationary processes from stationary
(2-14) processes via random transformations.

Model(19): Random periodic attenuation of stationary processes

Model(20): Multiplication of a stationary process by a cyclostationary process

c) Random time-scale transformations

i) Phase randomization:

Theorem(2-15): Reduction of cyclostationary processes to stationary processes via phase-randomization

Theorem(2-16): Equivalence between phase-randomizing and time-averaging of random processes.

Theorem(2-17): Preservation of system-cyclostationarity by phase-randomization.

Theorem(2-18): Reduction of cyclostationary systems to wide-sense-stationary systems via phase-randomization.

Theorem(2-19): Equivalence between phase-randomizing and time-averaging random systems.

ii) Generalized angle modulation:

Theorem(2-20): Preservation of cyclostationarity by a stationary time-scale transformation.

Theorem(2-21): Generation of cyclostationary processes from periodic signals via stationary time-scale transformations.

Model(21): Phase-modulated signal

Model(22): Frequency-modulated signal.

Theorem(2-22): Generation of cyclostationary processes from PAM time-scale transformations.

Model(23): Digital phase-modulation

Model(24): Digital frequency-modulation

Theorem(2-23): Generation of cyclostationary processes from stationary processes via cyclostationary time-scale transformations.

A random transformation, like any other random process, can be thought of as an ensemble of deterministic transformations. A single deterministic input to a random transformation yields an ensemble of deterministic outputs (a random output process) with each output resulting from one member of the ensemble of transformations. Of course, for any given event, only one member of the ensemble of transformations will occur and a single input will result in a single output. In this section we will be primarily concerned with random transformations of random processes. In particular, we will consider linear wide-sense-stationary and wide-sense-cyclostationary transformations (defined herein) in connection with the generation and preservation of cyclostationarity.

All random linear transformations of interest in this thesis will be integral transformations characterized by random impulse-response functions. Thus, input and output processes will be related as follows

$$x(t) = \int_{-\infty}^{\infty} g(t, \tau) y(\tau) d\tau \quad (2-29)$$

where $y(\tau)$ is a realization of the input process, $g(t, \tau)$ is a realization of the two-dimensional random impulse-response process, and $x(t)$ is the resultant realization of the output process. We will consider only those random transformations which are statistically independent of their inputs.

As pointed out by Zadeh [37], the system function (defined in Sec. 2) can be useful for characterizing random linear systems (transformations) as well as deterministic linear systems. Using the same notation as for deterministic systems, we denote the random system function for a random

system G as

$$\begin{aligned} H(t, f) &= \int_{-\infty}^{\infty} h(t, \tau) e^{-j2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} g(t, t-\tau) e^{-j2\pi f\tau} d\tau. \end{aligned}$$

In addition, we denote the autocorrelation function for the random system function as K_G and refer to it as the system autocorrelation function for G:

$$\begin{aligned} K_G(t, s, f, v) &\triangleq E\{H(t, f)H^*(s, v)\} \\ &= \iint_{-\infty}^{\infty} E\{h(t, \tau)h^*(s, \gamma)\} e^{-j2\pi(f\tau - v\gamma)} d\tau d\gamma. \end{aligned} \quad (2-30)$$

We note that K_G is a generalization¹⁵ of the quantity referred to as a system correlation function by Zadeh [37].

We proceed to derive an input-output relation for autocorrelation functions in terms of K_G . From Eq. (2-29) we have:

$$\begin{aligned} x(t)x^*(s) &= \iint_{-\infty}^{\infty} g(t, \tau)g^*(s, \gamma)y(\tau)y^*(\gamma)d\tau d\gamma \\ &= \iint_{-\infty}^{\infty} h(t, t-\tau)h^*(s, s-\gamma)y(\tau)y^*(\gamma)d\tau d\gamma \\ &= \iint_{-\infty}^{\infty} h(t, \tau)h^*(s, \gamma)y(t-\tau)y^*(s-\gamma)d\tau d\gamma. \end{aligned}$$

Equating the expected values of both sides of the above equations yields:

$$\begin{aligned} k_{xx}(t, s) &= \iint_{-\infty}^{\infty} E\{h(t, t-\tau)h^*(s, s-\gamma)\}k_{yy}(\tau, \gamma)d\tau d\gamma \\ &= \iint_{-\infty}^{\infty} E\{h(t, \tau)h^*(s, \gamma)\}k_{yy}(t-\tau, s-\gamma)d\tau d\gamma. \end{aligned} \quad (2-31)$$

¹⁵ $K_G(t, s, f, v)$ is also the inverse double Fourier transform w.r.t. f' and v' of $I_G(f', v', f, v)$ which is referred to as the bifrequency intensity spectrum and is the autocorrelation function for the bifrequency function for the system G (see, for example, Middleton [41]).

Using Fourier transforms in a manner similar to that used to derive Eq. (2-5) results in the equation:

$$k_{xx}(t,s) = \iint_{-\infty}^{\infty} K_G(t,s,f,v) k_{yy}(f,v) e^{j2\pi(ft-vs)} df dv. \quad (2-32)$$

Equations (2-30)-(2-32) will prove useful in the following development.

a) Linear wide-sense-stationary systems. An interesting subclass of random linear transformations are those which preserve wide-sense-stationarity: the output of a member of this class will be WSS if the input is WSS. Although the most general class of deterministic linear systems which preserves wide-sense-stationarity is the class of time-invariant filters, it is not true that the most general class of random linear systems which preserves wide-sense-stationarity is the class of random systems with ensemble members all of which are time-invariant filters. The following theorem--first stated and proved by Zadeh [37] for ergodic¹⁶ random processes and transformations--establishes a general class of random systems--to be called wide-sense-stationary (WSS) systems--which preserves wide-sense-stationarity:

THEOREM(2-10): If the system autocorrelation function for a random system G , whose input y is a zero-mean WSS process, satisfies the stationarity condition:

$$K_G(t,s,f,f) = K_G(t-s,0,f,f)$$

for all t,s,f then the output x is a zero-mean WSS process.

¹⁶ Zadeh used time averages rather than ensemble averages in his paper.

Proof:

i) From Eq. (2-29), the mean of the output x is

$$\begin{aligned} m_x(t) &= \int_{-\infty}^{\infty} E\{g(t, \tau)\} m_y(\tau) d\tau \\ &= 0 \quad \forall t \quad \text{since } m_y(\tau) = 0 \quad \forall \tau. \end{aligned}$$

ii) From Eq. (2-32), we have

$$\begin{aligned} k_{xx}(t, s) &= \iint_{-\infty}^{\infty} K_G(t, s, f, \nu) K_{yy}(f) \delta(f - \nu) e^{j2\pi(ft - \nu s)} df d\nu \\ &= \int_{-\infty}^{\infty} K_G(t, s, 0, f, f) K_{yy}(f) e^{j2\pi f(t - s)} df \\ &= k_{xx}(t - s) \quad \forall t, s \text{ (with some abuse of notation)} \end{aligned}$$

Hence, from i), ii) x is WSS.

QED

Notice that if all realizations of the random system G are time-invariant filters, then

$$\begin{aligned} K_G(t, s, f, f) &= E\{H(t, f)H^*(s, f)\} \\ &= E\{H(0, f)H^*(0, f)\} \\ &= K_G(0, 0, f, f) \quad \forall t, s, f \end{aligned}$$

so that G is clearly a WSS system.

Following are four specific examples of WSS systems which have been used in analyses dealing with statistical communication theory¹⁷ [38].

Notice that the realizations of the first WSS filter presented are time-

¹⁷A number of more sophisticated models for WSS systems (communication channels) are defined and discussed at length in the very "readable" report of Maurer [38]. Also, a number of models (for physical and biological applications in oceanography) for random time-varying channels are derived in the extensive four-part paper of Middleton [42].

varying systems (multipliers):

(1) Random WSS multipliers: If an input process y is multiplied by a WSS process z to yield an output process x , then realizations of the random impulse-response function which characterizes this system take the form

$$g(t, \tau) = z(t)\delta(t-\tau)$$

and $h(t, \tau) = g(t, t-\tau) = z(t)\delta(\tau)$

The system autocorrelation function is, from Eq. (2-30),

$$\begin{aligned} K_G(t, s, f, \nu) &= \iint_{-\infty}^{\infty} E\{z(t)\delta(\tau)z^*(s)\delta(\gamma)\}e^{-j2\pi(f\tau-\nu\gamma)}d\tau d\gamma \\ &= E\{z(t)z^*(s)\} \\ &= k_{zz}(t-s) \quad \forall t, s, f, \nu \end{aligned}$$

so that $K_G(t, s, f, \nu) = K_G(t-s, 0, f, \nu)$ for all t, s, f and G is a WSS system.

(2) Random first-order filter: The random first-order filter is simply a deterministic first-order filter with its inverse-time-constant replaced with a random variable ω so that we have

$$h(t, \tau) = e^{-2\pi\omega\tau}u(\tau)$$

where $u(\cdot)$ is the unit step function. Thus, if $p_\omega(\cdot)$ is the PDF for ω , then the inverse double Fourier transform of the system autocorrelation function K_G is

$$E\{h(t, \tau)h^*(s, \gamma)\} = \int_{-\infty}^{\infty} e^{-2\pi\sigma(\tau+\gamma)}p_\omega(\sigma)d\sigma u(\tau)u(\gamma)$$

(3) Random ideal low-pass filter: The random ideal low-pass filter is simply a deterministic ideal low-pass filter with its cutoff frequency replaced with the random variable ω so that

$$H(t, f) = \begin{cases} 1, & |f| \leq \omega \\ 0, & |f| > \omega \end{cases}$$

Thus, if $p_\omega(\cdot)$ is the PDF for ω , then

$$\begin{aligned} K_G(t, s, f, v) &= E\{H(t, f)H^*(s, v)\} \\ &= \int_{-\infty}^{\min(-|f|, -|v|)} p_\omega(\sigma) d\sigma + \int_{\max(|f|, |v|)}^{\infty} p_\omega(\sigma) d\sigma \\ &= 1 - \int_{-\max(|f|, |v|)}^{\max(|f|, |v|)} p_\omega(\sigma) d\sigma \\ &= 1 - P_\omega(\max(|f|, |v|)) + P_\omega(-\max(|f|, |v|)) \end{aligned}$$

where $P_\omega(\cdot)$ is the indefinite integral of $p_\omega(\cdot)$ and is, therefore, the probability distribution function for ω . Note that since all realizations of ω should be ≥ 0 , then the last term in the final expression is identically zero.

(4) Random delay: A frequently used model for a random communication channel is that of the random delay:

$$H(t, f) = e^{-j2\pi\Delta f}, \quad h(t, \tau) = \delta(\tau - \Delta)$$

where Δ is the random delay variable. If $p_\Delta(\cdot)$ is the PDF for Δ , then

$$\begin{aligned} K_G(t, s, f, v) &= \int_{-\infty}^{\infty} e^{-j2\pi\sigma(f-v)} p_\Delta(\sigma) d\sigma \\ &= P_\Delta(f-v), \end{aligned}$$

where $P_\Delta(\cdot)$ is the Fourier transform of $p_\Delta(\cdot)$ and is, therefore, the conjugate characteristic function for Δ .

The following class of CS signals are generated by the class of random time-invariant filters:

MODEL(18): Random periodic signals. In some biomedical experiments, a subject modeled as a random time-invariant linear system is excited by a known deterministic periodic signal. The measured response signal is then modeled as a random periodic signal: period known, but harmonic content random. In addition to these biomedical test signals, the periodic timing and test signals sent through communication transmission channels can also be modeled (at the receiver) as random periodic signals. All signals in this class are CS and take the form

$$x(t) = \sum_n a_n e^{j2\pi nt/T}$$

where $\{a_n\}$ is a random sequence with the constraint $a_n = a_{-n}^* \quad \forall n$.

It is apparent from the discussion and examples in Section 2 of this chapter that periodically varying systems are very common components in communication systems. Although a suitable deterministic model for a periodic system may well be known, it is not uncommon for the phase (time-origin) of the periodic variations to be unknown. In such cases it is sometimes sufficient to model the phase as a random variable uniformly distributed over one period. In fact, in some analyses, this is done even though the phase is known (this point will be discussed in Chapter IV). The following theorem establishes the fact that periodic systems with uniformly distributed random phases are WSS systems:

THEOREM(2-11): If the phase (time-origin) of a deterministic T-periodic system is modeled as a random variable θ uniformly distributed over one period, say $(-T/2, T/2)$, then the resultant random system is WSS.

Proof: The PDF for the uniformly distributed random variable θ is

$$P_{\theta}(\sigma) = \begin{cases} 1/T, & |\sigma| \leq T/2 \\ 0 & |\sigma| > T/2 \end{cases}.$$

Thus,

$$\begin{aligned} K_G(t, s, f, \nu) &= E\{H(t+\theta, f)H^*(s+\theta, \nu)\} \\ &= \int_{-\infty}^{\infty} H(t+\sigma, f)H^*(s+\sigma, \nu)P_{\theta}(\sigma)d\sigma \\ &= \frac{1}{T} \int_{-T/2}^{T/2} H(t+\sigma, f)H^*(s+\sigma, \nu)d\sigma \\ &= \frac{1}{T} \int_{s-T/2}^{s+T/2} H(t-s+\gamma, f)H^*(\gamma, f)d\gamma \end{aligned}$$

But since $H(t, f)$ is T -periodic in t , then the integrand is T -periodic in γ and the integral is simply an average over one period and is therefore independent of the absolute end points of the one-period interval, so that

$$\begin{aligned} K_G(t, s, f, \nu) &= \frac{1}{T} \int_{-T/2}^{T/2} H(t-s+\gamma, f)H^*(\gamma, f)d\gamma \\ &= \int_{-\infty}^{\infty} H(t-s+\gamma, f)H^*(\gamma, f)P_{\theta}(\gamma)d\gamma \\ &= K_G(t-s, 0, f, \nu) \quad \forall t, s, f, \nu. \end{aligned}$$

Hence by Theorem (2-10), G is a WSS system.

QED

i) Preservation of cyclostationarity. While it is true that all WSS systems preserve wide-sense-stationarity (by definition), it is not true--as might be expected--that all WSS systems preserve cyclostationarity. However, all the WSS system examples considered above do, indeed, preserve cyclostationarity. In fact, most WSS systems of practical interest do preserve cyclostationarity. This next theorem defines the most general subclass of WSS systems which preserve cyclostationarity:

THEOREM(2-12): If the system autocorrelation function for a random system G , whose input y is a zero-mean T-CS process, satisfies the stationarity condition:

$$K_G(t,s,f,v) = K(t-s, e^{j2\pi(ft-vs)}, f, v)$$

for all t,s,f,v and for some function K , then the output process x is a zero-mean T-CS process.

Proof: The mean of x is (as in the proof of Theorem (2-10)) obviously zero, and from Eq. (2-32) along with the stationarity condition we have

$$k_{xx}(t,s) = \iint_{-\infty}^{\infty} K(t-s, e^{j2\pi(ft-vs)}, f, v) k_{yy}(f, v) e^{j2\pi(ft-vs)} df dv.$$

Now, taking advantage of the periodicity of k_{yy} , we can expand its double Fourier transform K_{yy} into a series of the form of Eq. (2-14):

$$K_{yy}(f, v) = \sum_{n=-\infty}^{\infty} K_n(f) \delta(f-v+n/T) \quad \forall f, v,$$

so that

$$\begin{aligned} k_{xx}(t,s) &= \iint_{-\infty}^{\infty} K(t-s, e^{j2\pi(ft-vs)}, f, v) \sum_n K_n(f) \delta(f-v+n/T) \cdot \\ &\quad e^{j2\pi(ft-vs)} df dv \\ &= \sum_n \int_{-\infty}^{\infty} K(t-s, e^{j2\pi(f(t-s)-ns/T)}, f, f-n/T) K_n(f) \cdot \\ &\quad e^{j2\pi(f(t-s)-ns/T)} df \end{aligned}$$

Since this last expression is jointly T-periodic in t and s , then x is a zero-mean T-CS process.

QED

An example of a random system which satisfies the "stationarity condition" of both Theorems (2-10) and (2-12), but does not satisfy the more stringent and more obvious condition: $K_G(t,s,f,v) = K_G(t-s,0,f,v)$ for all t,s,f,v --which is satisfied by all previous examples--is the random S(2) delay. This random system has random impulse response

$$g(t,\tau) = \delta(\tau - z(t))$$

where z is any S(2) process, and its system function is

$$H(t,f) = e^{j2\pi f(t-z(t))}$$

so that its system autocorrelation function is

$$\begin{aligned} K_G(t,s,f,v) &= E\{e^{j2\pi(fz(t)-vz(s))}\} e^{-j2\pi(ft-vs)} \\ &= \int_{-\infty}^{\infty} e^{j2\pi(f\sigma-v\gamma)} p_{t-s}(\sigma,\gamma) d\sigma d\gamma e^{-j2\pi(ft-vs)} \\ &= P_{t-s}^*(f,v) e^{-j2\pi(ft-vs)} \end{aligned}$$

where $P_{t-s}(f,v)$ is the double Fourier transform w.r.t. σ and γ of $p_{t-s}(\sigma,\gamma)$ which is the joint PDF for $z(t)$ and $z(s)$. The realizations of the T-CS output x which result from any T-CS input y take the form

$$x(t) = y(z(t)).$$

Since z is S(2), then the time parameter for y does not increase monotonically and, in fact, "hovers" about the constant mean value of z , so that G is a rather peculiar transformation.

b) Linear cyclostationary systems. In the last subsection we defined the class of WSS systems to be those which, when driven with a

WSS input process, produce a WSS output process. We now define the class of cyclostationary (CS) systems to be those which, when driven with a WSS input, produce a CS output.

THEOREM(2-13): If the system autocorrelation function for a random system G , whose input y is a zero-mean WSS process, satisfies the cyclostationarity condition:

$$K_G(t,s,f,f) = K_G(t+T,s+T,f,f)$$

for all t,s,f and for some non-zero T , then the output process x is a zero-mean T -CS process.

Proof: From Eq. (2-32) we have

$$\begin{aligned} k_{xx}(t,s) &= \iint_{-\infty}^{\infty} K_G(t,s,f,v) K_{yy}(f) \delta(f-v) e^{j2\pi(ft-vs)} df dv \\ &= \int_{-\infty}^{\infty} K_G(t,s,f,f) K_{yy}(f) e^{j2\pi f(t-s)} df \\ &= \int_{-\infty}^{\infty} K_G(t+T,s+T,f,f) K_{yy}(f) e^{j2\pi f((t+T)-(s+T))} df \\ &= k_{xx}(t+T,s+T) \quad \forall t,s, \end{aligned}$$

and the mean for x is clearly zero. Hence, x is T -CS.

QED

A large subclass of CS systems is the class of random periodic systems. The members of this class have ensembles with elements all of which are periodic deterministic systems. These random periodic systems can be characterized by a random system function H which is a one-dimensional random process indexed by f with parameter t , such that the random processes $H(t+T, \cdot)$ and $H(t, \cdot)$ are identical for all t .

Such random system functions can be represented by a random Fourier series

$$H(t, \cdot) = \sum_{n=-\infty}^{\infty} H_n(\cdot) e^{j2\pi n t/T}$$

where the coefficients are random functions. Thus, paralleling deterministic periodic systems, we see that random periodic systems have realizations with the structure shown in Figure (2-2) where the path-filter transfer functions are realizations of the random coefficient functions H_n .

The fact that random periodic systems are CS systems is obvious, but will be stated here as a theorem for convenient reference:

THEOREM(2-14): If the input to a random T-periodic system is WSS, then the output is T-CS.

This theorem is a generalization of Theorem (2-4) which applies to deterministic periodic systems and serves as a basis for the generation of many CS processes of interest. This more general theorem (2-14) also serves as a basis for the generation of CS processes, and to illustrate this we give two examples of random periodic systems:

MODEL(19): Random periodic attenuation of a WSS process. The random periodic attenuator is a useful model for some communication channels and for some modulators, and has a random impulse-response function with realizations of the form

$$g(t, \tau) = p(t) \delta(t - \tau),$$

where p is a random periodic function. Thus, the random system function takes the form

$$\begin{aligned}
 H(t, f) &= \int_{-\infty}^{\infty} p(t) \delta(\tau) e^{-j2\pi f \tau} d\tau \\
 &= \sum_n p_n e^{j2\pi n t / T}
 \end{aligned}$$

where $\{p_n\}$ are the random Fourier coefficients for p .

Another useful model for communication channels is a deterministic periodic attenuation followed by a random dispersion. The random impulse-response function for this model has realizations of the form

$$g(t, \tau) = p(\tau) g_0(t - \tau)$$

where $p(t)$ is the periodic attenuator and g_0 is the random time-invariant impulse-response function. Thus,

$$\begin{aligned}
 H(t, f) &= \int_{-\infty}^{\infty} p(t - \tau) g_0(\tau) e^{-j2\pi f \tau} d\tau \\
 &= \sum_n G_0(f + n/T) p_n e^{j2\pi n t / T}
 \end{aligned}$$

where $\{p_n\}$ are the deterministic Fourier coefficients for $p(t)$ and G_0 is the random transfer function for the dispersion.

The following is an example of a CS system which is not a periodic system:

MODEL(20): Multiplication of a WSS process by a cyclostationary process.

In communication systems, signals are sometimes subjected to multiplicative noise. If the noise is a CS process, then we have a CS system with impulse response realizations of the form

$$g(t, \tau) = n(t) \delta(t - \tau)$$

where n is a CS process, and the system autocorrelation function is

$$K_G(t, s, f, \nu) = k_{nn}(t, s).$$

c) Random time-scale transformations. In Section 2 we discussed deterministic time-scale transformations of the form $t'(t) = t + p(t)$ where $p(t)$ was a deterministic periodic function. Here, we will consider random time-scale transformations with realizations of the form $t'(t) = t + z(t)$ where z is a random process which is stationary (or, in some cases, cyclostationary) of order two ($S(2)$).

i) Phase randomization. The simplest possible random time-scale transformation is the random phase-shift: $t'(t) = t + \theta$, where θ is a single random variable with PDF $p_\theta(\cdot)$. The phase-shift, as defined here, is identical to the random delay which was discussed in subsection 5a. Thus, we know by Theorem (2-12) that random phase shifts preserve cyclostationarity. That is, it is not necessary that the time-origin or phase of periodicity be known precisely (be deterministic) in order that a process be CS. In fact, the phase can be a random variable with arbitrary PDF and the process still be CS. Furthermore, there is an interesting class of PDF's for the random phase which reduce all CS processes to stationary processes. This result, which is similar to that of Theorem (2-11) is established in the following theorem;¹⁸

THEOREM(2-15): For every T-CS random process y , the random process x with realizations of the form $x(t) = y(t+\sigma)$ is WSS if and only if the PDF $p_\theta(\cdot)$ for the independent random-phase variable, with realization σ , satisfies the condition:

¹⁸This theorem (in a slightly different form) has been stated and proved by Hurd [5], and the "sufficient" portion has been proved and/or demonstrated for specific processes by a number of authors [1,2,3,7]. In addition, Hurd has given an interesting companion theorem which states that a process is CS if and only if its phase-randomized version is WSS.

$$\sum_{n=-\infty}^{\infty} p_{\theta}(\sigma - nT) = 1/T \quad \forall \sigma;$$

and if the PDF is constrained to be duration limited to $(-1/2T, 1/2T)$, then this necessary and sufficient condition reduces to

$$p_{\theta}(\sigma) = \begin{cases} 1/T, & |\sigma| \leq 1/2T \\ 0, & |\sigma| > 1/2T \end{cases}$$

Proof:

i) The mean for x is

$$m_x(t) = E\{y(t+\theta)\} = \int_{-\infty}^{\infty} m_y(t+\sigma) p_{\theta}(\sigma) d\sigma.$$

Now, since m_y is T -periodic, we make use of its Fourier series representation

$$\begin{aligned} m_x(t) &= \int_{-\infty}^{\infty} \sum_n a_n e^{j2\pi n(t+\sigma)/T} p_{\theta}(\sigma) d\sigma \\ &= \sum_n a_n P_{\theta}^*(n/T) e^{j2\pi nt/T} \end{aligned}$$

where $P_{\theta}(\cdot)$ is the Fourier transform of $p_{\theta}(\cdot)$. Clearly, from this last expression, m_x will be independent of t for every sequence $\{a_n\}$ if and only if $P_{\theta}(n/T) = 0$ for all $n \neq 0$.

ii) The autocorrelation for x is

$$k_{xx}(t,s) = E\{y(t+\theta)y(s+\theta)\} = \int_{-\infty}^{\infty} k_{yy}(t+\sigma, s+\sigma) p_{\theta}(\sigma) d\sigma.$$

Now, since $k_{yy}(t,s)$ is jointly T -periodic in t and s , we make use of its Fourier series representation (see Eq. (2-13)):

$$\begin{aligned} k_{xx}(t,s) &= \int_{-\infty}^{\infty} \sum_n k_n(t-s) e^{j2\pi n(t+\sigma)/T} p_{\theta}(\sigma) d\sigma \\ &= \sum_n k_n(t-s) P_{\theta}^*(n/T) e^{j2\pi nt/T} \end{aligned}$$

Clearly, from this last expression, $k_{xx}(t,s)$ will be a function of only $(t-s)$ for every sequence $\{k_n\}$ if and only if $p_\theta(n/T) = 0$ for every $n \neq 0$.

iii) From the Poisson sum formula [2]:

$$\begin{aligned}\sum_m p_\theta(\sigma - m/T) &= \frac{1}{T} \sum_n p_\theta(n/T) e^{-j2\pi n\sigma/T} \\ &= \frac{1}{T} p_\theta(0) = \frac{1}{T} \int_{-\infty}^{\infty} p_\theta(\gamma) d\gamma = 1/T \quad \forall \sigma,\end{aligned}$$

where the second line is a valid equality if and only if $p_\theta(n/T) = 0$ for all $n \neq 0$. Hence, from i), ii), iii), x is WSS if and only if

$$\sum_{n=-\infty}^{\infty} p_\theta(\sigma - nT) = 1/T \quad \forall \sigma$$

and clearly, if $p_\theta(\cdot)$ is constrained to be duration limited to $(-1/2T, 1/2T)$, then this condition can only be satisfied by the uniform PDF

$$p_\theta(\sigma) = \begin{cases} 1/T, & |\sigma| \leq 1/2T \\ 0, & |\sigma| > 1/2T. \end{cases}$$

QED

This next theorem establishes a rather obvious, but quite useful, result on the relationship between the statistics of a phase-randomized process and the time-averaged (smoothed) statistics of its deterministic-phase counterpart.

THEOREM(2-16): The first and second order statistics of the process x obtained by phase-randomizing a process y , with a random-phase PDF $p_\theta(\cdot)$, are identical to the time-averaged (smoothed) statistics of y if the weighting function used for averaging is $p_\theta(\cdot)$.

Proof: For any function $f(\cdot)$ we have

$$\begin{aligned} E\{f(x(t+\theta))\} &= \int_{-\infty}^{\infty} E\{f(x(t+\sigma))\} p_{\theta}(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} E\{f(x(t+\sigma))\} p_{\theta}(\sigma) d\sigma / \int_{-\infty}^{\infty} p_{\theta}(\sigma) d\sigma, \end{aligned}$$

and this last expression is the time-averaged value of $E\{f(x(t))\}$ with weighting function $p_{\theta}(\cdot)$; and for any function $g(\cdot, \cdot)$ we have

$$\begin{aligned} E\{g(x(t+\theta), x(s+\theta))\} &= \int_{-\infty}^{\infty} E\{g(x(t+\sigma), x(s+\sigma))\} p_{\theta}(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} E\{g(x(t+\sigma), x(s+\sigma))\} p_{\theta}(\sigma) d\sigma / \int_{-\infty}^{\infty} p_{\theta}(\sigma) d\sigma \end{aligned}$$

and this last expression is the time-averaged value of $E\{g(x(t), x(s))\}$ with weighting function $p_{\theta}(\cdot)$.

QED

A direct result of this theorem is the fact that the mean and autocorrelation functions for the WSS process x obtained by time-averaging over one period the mean and autocorrelation functions for any T-CS process y :

$$\begin{aligned} m_x &= \frac{1}{T} \int_{-T/2}^{T/2} m_y(t) dt \\ k_{xx}(t-s) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{yy}(t+\sigma, s+\sigma) d\sigma, \end{aligned}$$

are identical to the mean and autocorrelation functions for the WSS process x' obtained by phase-randomizing y with a random phase variable uniformly distributed over one period $(-1/2T, 1/2T)$.

The three preceding results generalize from one-dimensional random signal processes to two-dimensional random system processes as follows:

THEOREM(2-17): For every T-CS system G' with random system function $H'(t,f)$, the random system G with random system function realizations of the form $H(t,f) = H'(t+\sigma,f)$ where σ is the realization of an independent random-phase variable with arbitrary PDF, is also a T-CS system.

THEOREM(2-18): For every T-CS system G' , the corresponding phase-randomized system G is WSS if and only if the PDF $p_\theta(\cdot)$ for the independent random-phase variable satisfies the condition

$$\sum_{n=-\infty}^{\infty} p_\theta(\sigma-nT) = 1/T \quad \forall \sigma,$$

and if the PDF is constrained to be duration limited to $(-1/2T, 1/2T)$, then this necessary and sufficient condition becomes

$$p_\theta(\sigma) = \begin{cases} 1/T, & |\sigma| \leq 1/2T \\ 0, & |\sigma| > 1/2T. \end{cases}$$

THEOREM(2-19): The system autocorrelation function for the random system G obtained by phase-randomizing a system G' , with a random-phase PDF $p_\theta(\cdot)$ is identical to the time-averaged (smoothed) system autocorrelation function for G' , if the weighting function used for averaging is $p_\theta(\cdot)$.

The proofs of these last two theorems directly parallel those for Theorems (2-15), (2-16), and will not be repeated here. The proof of Theorem (2-17) is trivial.

As a result of Theorems (2-11), (2-16), (2-19), we see that when it is appropriate in specific analyses to ignore periodic fluctuations in random cyclostationary signals and systems and in deterministic periodic signals and systems, then these fluctuations can be "averaged out" either by uniformly randomizing, over one period, the time-origin or phase of the

periodic fluctuations, or by time-averaging, over one period, the fluctuations with a uniform weighting function.

Also, from the discussion at the beginning of this subsection and Theorem (2-17), we see that it is not necessary that the time-origin or phase of periodicity be known exactly (be deterministic) in order that a signal or system be cyclostationary. In fact, the phase can be a random variable with arbitrary PDF.

ii) Generalized angle modulation. We now consider more general random time-scale transformations.

THEOREM(2-20): A random time-scale transformation G , with realizations of the form $t'(t) = t + z(t)$ where z is a $S(2)$ process with joint PDF for $z(t)$ and $z(s)$ denoted $p_{t-s}(\cdot, \cdot)$, has system autocorrelation function

$$K_G(t, s, f, \nu) = P_{t-s}^*(f, \nu)$$

where P_{t-s} is the double Fourier transform of p_{t-s} , and the transformation G preserves cyclostationarity.

Proof: The impulse-response function for G has realizations of the form

$$g(t, \tau) = \delta(\tau - t - z(t))$$

so that

$$h(t, \tau) = \delta(\tau + z(t))$$

and

$$H(t, f) = e^{j2\pi z(t)f}$$

Now, the system autocorrelation function is

$$\begin{aligned} K_G(t, s, f, \nu) &= \iint_{-\infty}^{\infty} e^{j2\pi f\sigma} e^{-j2\pi \nu \gamma} p_{t-s}(\sigma, \gamma) d\sigma d\gamma \\ &= P_{t-s}^*(f, \nu). \end{aligned}$$

This system autocorrelation function clearly satisfies the hypothesis of Theorem (2-12); hence, the system G preserves the cyclostationarity of zero-mean processes. But since the mean of the output x satisfies

$$\begin{aligned} m_x(t) &= \int_{-\infty}^{\infty} m_y(t+\sigma)p(\sigma)d\sigma \\ &= \int_{-\infty}^{\infty} m_y(t+T+\sigma)p(\sigma)d\sigma \\ &= m_x(t+T). \end{aligned}$$

where y is the input to G and $p(\cdot)$ is the PDF for $z(t)$ for all t , then G also preserves cyclostationarity of non-zero-mean processes.

QED

The degenerate case in which the input process is a deterministic T -periodic signal is actually generalized angle modulation, and is of special interest in communications. For convenient reference, we restate the above theorem for this special case:

THEOREM(2-21): If a deterministic T -periodic signal $q(t)$ is subjected to a random time-scale modulation with realizations of the form $t'(t)=t + z(t)$ where z is $S(2)$, then the resultant process x with realizations of the form $q(t + z(t))$ is T -CS.

Examples:

MODEL(21): Conventional phase-modulation (PM). If $q(t)$ is a sinusoid, then x is a conventional phase-modulated signal where z is the modulating signal.

MODEL(22): Conventional frequency-modulation (FM). If $q(t)$ is a sinusoid, and if the realizations of z take the form

$$z(t) = \int_{-\infty}^t z_0(\sigma) d\sigma$$

then x is a conventional frequency modulated signal where z_0 is the modulating signal.

Variations on these two commonly used communication signal formats are those of digital PM and digital FM, the statistics of which have been partially analyzed by various authors [23], [24]. The following theorem establishes the cyclostationarity of digital angle-modulated signals under a stationarity assumption which differs from that in the above theorem.

THEOREM(2-22): If a deterministic T -periodic signal $q(t)$ is subjected to a random time-scale modulation with realizations of the form $t'(t) = t + z(t)$ where z is a PAM signal with realizations of the form

$$z(t) = \sum_n a_n p_0(t-nLT)$$

where $\{a_n\}$ is a $S(2)$ random sequence and L is an integer, then the resultant signal x with realizations of the form

$$x(t) = q(t + \sum_n a_n p_0(t-nLT))$$

is CS with period LT .

Proof:

$$\begin{aligned} \text{i)} \quad m_x(t) &= \int_{-\infty}^{\infty} q(t + \sum_n a_n p_0(t-nLT)) p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} q(t + LT + \sigma \sum_n p_0(t + LT - (n+1)LT)) p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} q(t + LT + \sigma \sum_{n'} p_0(t + LT - n'LT)) p(\sigma) d\sigma \\ &= m_x(t + LT) \quad \forall t, \end{aligned}$$

where $p(\cdot)$ is the PDF for a_n for all n ,

ii)

$$\begin{aligned}
 k_{xx}(t,s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(t + \sum_n \sigma p_0(t-nLT)) q(s + \sum_m \gamma p_0(s-mLT)) p_{n-m}(\sigma, \gamma) d\sigma d\gamma \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(t + LT + \sigma \sum_n p_0(t + LT - (n+1)LT)) \cdot \\
 &\quad q(s + \gamma \sum_m p_0(s + LT - (m+1)LT)) p_{(n+1)-(m+1)}(\sigma, \gamma) d\sigma d\gamma \\
 &= k_{xx}(t + LT, s + LT) \quad \forall t, s,
 \end{aligned}$$

where $p_{n-m}(\cdot, \cdot)$ is the joint PDF for a_n and a_m for all n and m . Hence, from i) and ii), x is CS with period LT .

QED

Examples:

MODEL(23): Digital phase-modulation (DPM). If $q(t)$ is a sinusoid, then x is a conventional digital phase-modulated signal with modulating sequences $\{a_n\}$, and pulse-shape $p_0(t)$.

MODEL(24): Digital frequency-modulation (DFM). If $q(t)$ is a sinusoid, and if the "pulses" $p_0(t)$ take the form

$$p_0(t) = \int_{-\infty}^t p'_0(\sigma) d\sigma,$$

then x is a conventional digital frequency-modulated signal with modulating sequence $\{a_n\}$, and pulse-shape $p'_0(t)$.

Finally, we point out that random time-scale transformations can be CS systems in that they can transform WSS inputs into CS outputs:

THEOREM(2-23): The output x of a random time-scale transformation G with realizations of the form $t'(t) = t + z(t)$, where z is a T-CS(2) process, is T-CS if the input y is WSS.

Proof: The system autocorrelation function for G is (as shown in the proof of Theorem (2-20))

$$K_G(t, s, f, v) = P_{t, s}^*(f, v),$$

where $P_{t, s}$ is the double Fourier transform of $p_{t, s}$ which is the joint PDF for $z(t)$ and $z(s)$. But since z is T-CS(2) then $p_{t+T, s+T} = p_{t, s}$ for all t and s , so that K_G satisfies the hypothesis of Theorem (2-13), and G preserves cyclostationarity of zero-mean inputs. But since the mean for the output x takes the form

$$\begin{aligned} m_x(t) &= \int_{-\infty}^{\infty} m_y p_t(\sigma) d\sigma = m_y \quad \forall t \\ &= m_x(t + T) \end{aligned}$$

where p_t is the PDF for $z(t)$, then G also preserves the cyclostationarity of non-zero-mean inputs y .

QED

6. Random Multi-dimensional Linear Transformations (Random Scanning)

In this section, we briefly present a generalization of the ideas presented in Section 3 on deterministic multi-dimensional linear transformations. The motivation here is to develop a theory upon which studies of random scanners can be based. We use the same definitions and input-output relation (Eq. (2-21)) as in Section 3; however, we now allow the three-dimensional impulse-response function to be a (four-dimensional) random process. Thus, generalizing our definition of "quasi-periodic transformation" (Section 3, Eq. (2-22)) and our definition of "cyclostationary system" (Section 5b), we define the transformation G to be quasi-cyclostationary with period T if and only if the generalized system autocorrelation function K_G satisfies the cyclostationarity condition:

$$K_G(t+T, s+T, f, f, u_1, v_1, u_2, v_2) = K_G(t, s, f, f, u_1, v_1, u_2, v_2)$$

for all $t, s, f, u_1, v_1, u_2, v_2$ where K_G is defined as in Eq. (2-30) of Section 5. (In most applications u_1, u_2, v_1, v_2 will be position coordinates; t, s will be time variables, and f will be the frequency variable.)

Now, we have the following generalization of Theorems (2-6), (2-13):

THEOREM(2-24): If the input y to a three-dimensional quasi-cyclostationary transformation with period T is a zero-mean quasi-wide-sense-stationary process, then the one-dimensional output process x is T-CS.

The proof of this theorem is a simple extension of that for Theorem (2-13).

For applications of this theorem, we have the generalizations of those applications of Theorem (2-6) in Section 3: i.e., random scanners. Consider, for example, the line scanner of Section 3, but with the random window function

$$g(t, \tau, u, v) = g(u_1(t) - u, v_1(t) - v) \delta(t - \tau),$$

where g is a two-dimensional random function, and v_1 and u_1 are one-dimensional random processes. If g is statistically independent of u_1 and v_1 which are either jointly CS(4) (cyclostationary of order 4) or are independent and CS(2), then the random line-scanner G satisfies the hypotheses of Theorem (2-24). Hence, random line-scanners with random (randomly selected) window shapes and random CS window velocities generate CS signals from WSS inputs.

Thus, we see that the precise (deterministic) periodicity used in the video line-scanning model of Section 3 is not necessary in order for the resultant video signal to be modeled as CS.

More generally, we see that signals obtained from various scanning operations can be modeled as cyclostationary even if there is no precise (deterministic) periodicity in the model of the scanner.

7. Nonlinear Random Transformations (Jitter)

Unlike the random linear transformations of the preceding sections, there does not exist one simple characterization, such as a system function, for the random nonlinear transformations of interest in this thesis. The periodic Volterra series representation introduced in Section 4 on deterministic nonlinear transformations could be generalized here by allowing the kernel functions to be random; however, such a representation would not serve the purposes of this thesis.

In this section, then, we briefly extend the results on deterministic zero-memory nonlinearities presented in Section 4, and then turn to the important issue of jitter. We characterize this random timing disturbance in general terms, and then show that it preserves cyclostationarity under a fairly liberal set of conditions.

a) Zero-memory random nonlinearities. A nonlinear random transformation is said to have zero-memory if and only if for every member of its ensemble of transformations, every image function evaluated at any time, say t_0 , depends on the corresponding inverse image function evaluated only at the same time t_0 . Realizations of the input and output of such transformations are related as follows

$$x(t) = G(y(t), t)$$

where $G(\cdot, \cdot)$ is a realization of the random transformation G .

We begin with a theorem which establishes a class of such transformations which generate CS processes from $S(2)$ processes:

THEOREM(2-2.5): If the input y to a zero-memory random nonlinearity G is $S(2)$, and if G satisfies the cyclostationarity condition:

$$E\{G(\sigma, t)\} = E\{G(\sigma, t+T)\} \quad \forall t, \sigma$$

$$E\{G(\sigma, t)G(\gamma, s)\} = E\{G(\sigma, t+T)G(\gamma, s+T)\} \quad \forall t, s, \sigma, \gamma,$$

then the output x is T -CS.

Proof:

$$\begin{aligned} m_x(t) &= E\{G(y(t), t)\} \\ &= \int_{-\infty}^{\infty} E\{G(\sigma, t)\} p(\sigma) d\sigma \\ &= \int_{-\infty}^{\infty} E\{G(\sigma, t+T)\} p(\sigma) d\sigma \\ &= m_x(t+T), \end{aligned}$$

where $p(\cdot)$ is the PDF for $y(t)$ for all t , and

$$\begin{aligned} k_{xx}(t, s) &= E\{G(y(t), t)G(y(s), s)\} \\ &= \iiint_{-\infty}^{\infty} E\{G(\sigma, t)G(\gamma, s)\} p_{t-s}(\sigma, \gamma) d\sigma d\gamma \\ &= \iiint_{-\infty}^{\infty} E\{G(\sigma, t+T)G(\gamma, s+T)\} p_{(t+T)-(s+T)}(\sigma, \gamma) d\sigma d\gamma \\ &= k_{xx}(t+T, s+T) \quad \forall t, s, \end{aligned}$$

where $p_{t-s}(\cdot, \cdot)$ is the joint PDF for $y(t)$ and $y(s)$. Hence x is T -CS.

QED

As an example of a cyclostationary zero-memory nonlinearity, consider the very elementary transformation of "additive noise":

$G(y(t), t) = y(t) + n(t)$, where n is a CS random "noise" process. G clearly

satisfies the cyclostationarity condition of the theorem. In fact, the sum of any number of statistically independent WSS and T-CS processes will be T-CS.

b) Jitter. The various repetitive operations such as sampling, scanning, and multiplexing, to which signals are often subjected, are-- of course--never exactly periodic. In some analyses, this departure from exact periodicity can be neglected; however, in the event that it is-- or may be--significant, this departure from exact periodicity should be modeled so that it can be dealt with analytically.

We have seen, throughout this chapter, a number of models for T-CS processes with realizations of the form

$$y(t) = \sum_{n=-\infty}^{\infty} x_n(t-nT)$$

where, for example, the translates $\{x_n(t-nT)\}$ might be one-line segments from a line-scanned rectangular field as in Video, or finite length records from a multitude of sources as in time-division multiplexing schemes, or individual random pulses as in the various synchronous pulse modulation schemes such as PAM, FSK, and others. For CS processes with realizations of this form, we model departures from exact periodicity by incorporating a random epoch-jitter variable δ_n in each translate so that

$$x(t) = \sum_n x_n(t-nT+\delta_n)$$

where $\{\delta_n\}$ is a random sequence. Hence x is obtained from y via a random nonlinear transformation with memory.

Now, two fundamental models for this sequence $\{\delta_n\}$ are of importance, in that each accurately represents one of the two most common physical

mechanisms which cause jitter:

(1) each translate is jittered w.r.t. the time-origin of the total process ($t = 0$),

(2) each translate, say the n^{th} , is jittered w.r.t. the time-origin of the preceding translate ($t = (n-1)T + \sum_{i=0}^{n-1} \delta_i$).

In the first case, where "jitter" is the more appropriate terminology, the random sequence $\{\delta_n\}$ can be modeled as a stationary finite-variance sequence with autocorrelation function denoted $k_{\delta\delta}(n-m)$. Then the second case, for which the term jitter is somewhat inappropriate, can be modeled by a new sequence $\{\gamma_n\}$ which is composed of the partial sums of the $\{\delta_n\}$:

$$x(t) = \sum_n x_n(t - nT + \gamma_n)$$

where

$$\gamma_n = \sum_{i=0}^n \delta_i.$$

As an example consider the case where the $\{\delta_n\}$ are all independent with zero mean and equal variance σ^2 . Then

$$k_{\delta\delta}(n-m) = \begin{cases} \sigma^2, & n = m \\ 0, & n \neq m \end{cases},$$

$$\begin{aligned} k_{\gamma\gamma}(n,m) &= E\{\gamma_n \gamma_m\} \\ &= \sum_{i=0}^n \sum_{j=0}^m E\{\delta_i \delta_j\} \\ &= \sum_{j=0}^{\min(n,m)} E\{\delta_j^2\} \\ &= \sigma^2(\min(n,m)+1), \quad n, m \geq 0. \end{aligned}$$

Thus, $\{\gamma_n\}$ is a discrete Wiener process (Wiener sequence) and is non-stationary with growing variance. Due to the difficulties associated with

this model, we will simply state that the corresponding jittered process x is not, in general, CS; but, in the event that the process is observed by a tracking receiver, then the observed process can be remodeled--using the time-varying frame of reference provided by the tracker--in terms of the first model for jitter (mentioned above) which is more tractable.

The new text book Theory of Synchronous Communications by J.J. Stifler [40] provides a comprehensive tutorial treatment of synchronization: the problems of detecting and estimating symbol epochs and carrier phase, and of tracking fluctuations in periodicity, and related synchronization problems are treated using the theories of maximum-likelihood detection and estimation, and familiar communications components such as matched filters or correlators and phase-lock loops.

Returning now to the first model for jitter, we establish the preservation of cyclostationarity with the following theorem:

THEOREM(2-26): If a T-CS process y with realizations of the form

$$y(t) = \sum_{n=-\infty}^{\infty} x_n(t-nT)$$

and autocorrelation function of the form

$$k_{yy}(t,s) = \sum_{n,m} k_{n-m}(t-nT, s-mT)$$

is modified to include jitter:

$$x(t) = \sum_n x_n(t-nT+\delta_n)$$

where the random jitter sequence $\{\delta_n\}$ is S(2), then the jittered process x is T-CS.

Proof:

$$\begin{aligned}
 \text{i)} \quad m_x(t) &= \sum_n \int_{-\infty}^{\infty} E\{x_n(t-nT+\sigma)\} p(\sigma) d\sigma_n \\
 &= \int_{-\infty}^{\infty} E\left\{\sum_n x_n(t-nT+\sigma)\right\} p(\sigma) d\sigma \\
 &= \int_{-\infty}^{\infty} E\{y(t+\sigma)\} p(\sigma) d\sigma \\
 &= \int_{-\infty}^{\infty} m_y(t+\sigma) p(\sigma) d\sigma \\
 &= \int_{-\infty}^{\infty} m_y(t+T+\sigma) p(\sigma) d\sigma \\
 &= m_x(t+T) \quad \forall t,
 \end{aligned}$$

where $p(\cdot)$ is the PDF for δ_n for every n .

ii)

$$\begin{aligned}
 k_{xx}(t,s) &= \sum_{n,m} \iint_{-\infty}^{\infty} E\{x_n(t-nT+\sigma) x_m(s-mT+\gamma)\} p_{n-m}(\sigma, \gamma) d\sigma d\gamma \\
 &= \iint_{-\infty}^{\infty} E\left\{\sum_{n,m} x_n(t-nT+\sigma) x_m(s-mT+\gamma)\right\} p_{n-m}(\sigma, \gamma) d\sigma d\gamma \\
 &= \iint_{-\infty}^{\infty} \sum_{n,m} k_{n-m}(t-nT+\sigma, s-mT+\gamma) p_{n-m}(\sigma, \gamma) d\sigma d\gamma \\
 &= \iint_{-\infty}^{\infty} \sum_{n,m} k_{(n+1)-(m+1)}(t+T-(n+1)T+\sigma, s+T-(m+1)T+\gamma) \cdot \\
 &\quad p_{(n+1)-(m+1)}(\sigma, \gamma) d\sigma d\gamma \\
 &= \iint_{-\infty}^{\infty} \sum_{n',m'} k_{n'-m'}(t+T-n'T+\sigma, s+T-m'T+\gamma) p_{n'-m'}(\sigma, \gamma) d\sigma d\gamma \\
 &= k_{xx}(t+T, s+T) \quad \forall t, s,
 \end{aligned}$$

where $p_{n-m}(\cdot, \cdot)$ is the joint PDF for δ_n and δ_m for all n and m . Now, from i) and ii), x is T-CS.

QED

As examples, we merely note that the hypotheses of this theorem are satisfied by the time-division-multiplexing scheme (TDM₁) with independent sources, by the video model derived by Franks (see Sec. 3), and by all the synchronous pulse-modulation schemes, with S(2) modulating sequences, such as PAM, FSK, PWM, and others (see Sec. 4).

We now give an extension of this theorem to CS random systems:

THEOREM(2-27): If a T-CS random system G', with system function realizations of the form

$$H'(t, f) = \sum_{n=-\infty}^{\infty} H_n(t-nT, f)$$

and system autocorrelation function of the form

$$K_{G'}(t, s, f, \nu) = \sum_{n, m} K_{n-m}(t-nT, s-mT, f, \nu)$$

is modified to include jitter

$$H(t, f) = \sum_n H_n(t-nT+\delta_n, f)$$

where $\{\delta_n\}$ is S(2), then the resultant jittered system G is T-CS.

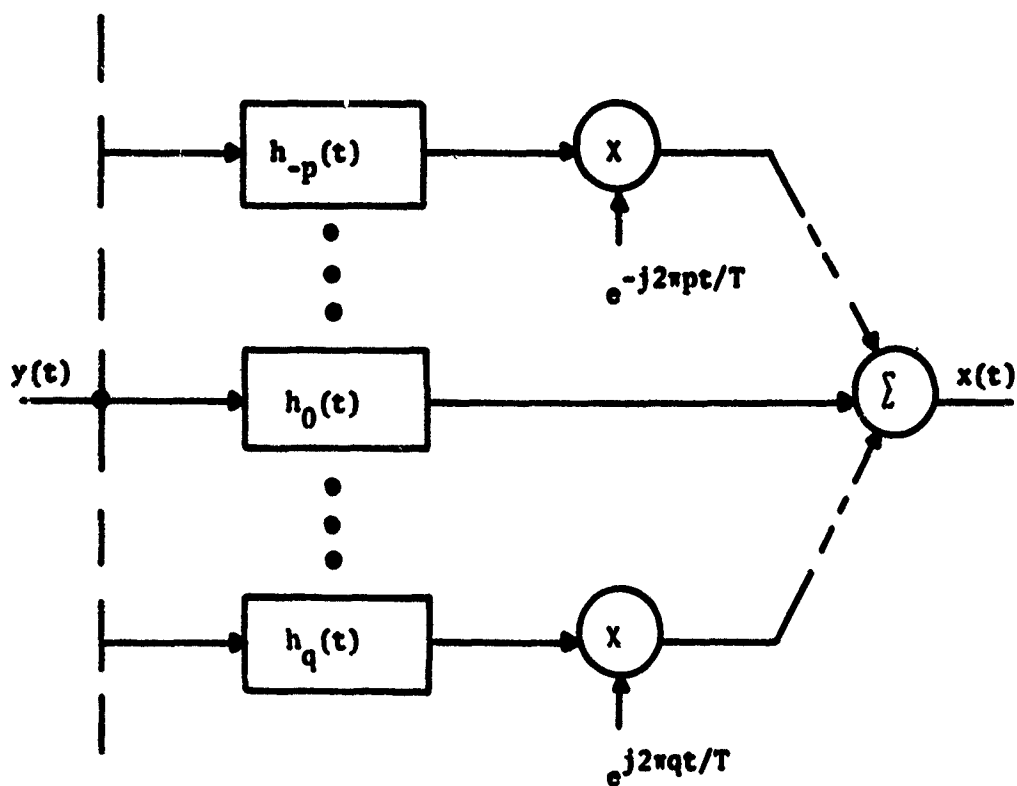
The proof directly parallels part ii) for Theorem (2-26) and will not be repeated here. Note that, for the special case of deterministic signals and systems, the two preceding theorems establish the fact that T-periodic signals and systems, when jittered, become T-CS signals and systems.

In conclusion, we emphasize the fact that there need be no exact periodicity in the realizations of a CS process: the occurrence of pulses in jittered PAM, for example, is not periodic, nor are the pulse occurrences in a Poisson pulse process with periodic rate parameter.

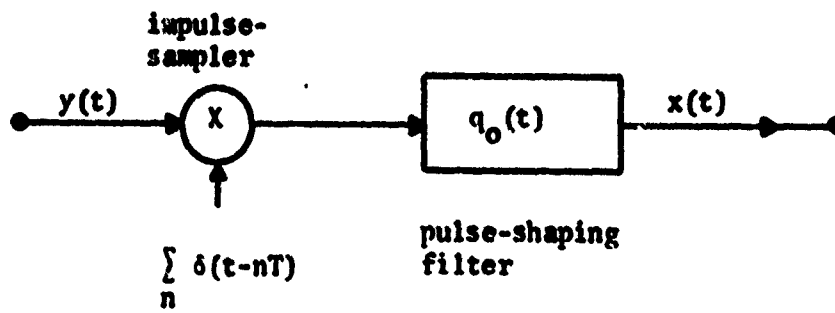
The periodicity need only be present in the first and second order ensemble averages. Yet, as pointed out a number of times in both this chapter and Chapter I, the mean and variance for a CS process can be constant so that the periodicity in these quantities is degenerate, and still such processes can exhibit strong periodic fluctuations in correlation, as exemplified by the time-division-multiplex (TDM_1 or TDM_2) of WSS processes with equal means and variances but with disjoint PSD's, or as illustrated in Figure (1-3) for PAM.



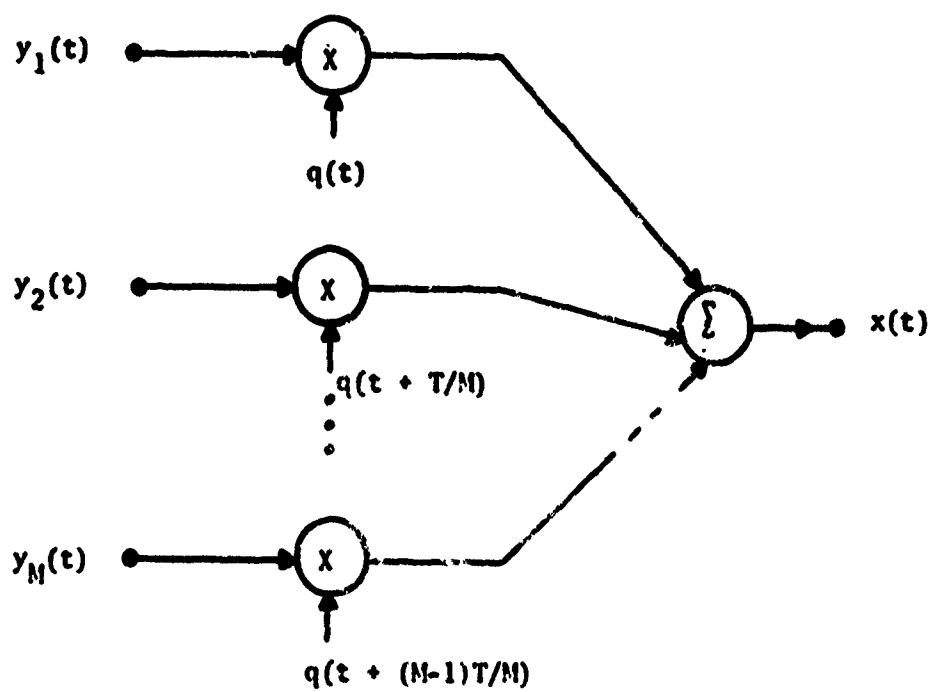
Figure(2-1) Diagram of system(transformation) G with input (inverse-image) y and output(image) x .



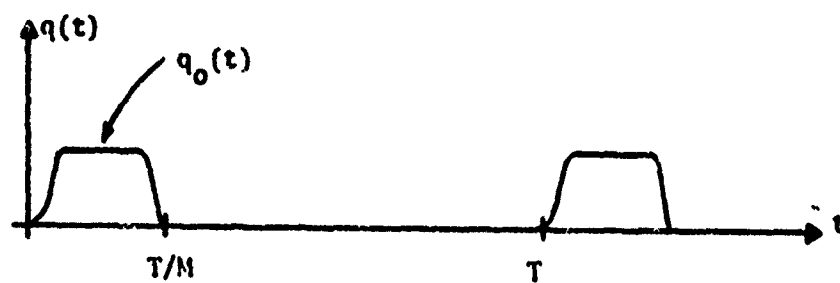
Figure(2-2) Structural diagram of the Fourier series representation of a periodically(T) time-varying linear system.



Figure(2-3) Generation of PAN signal.

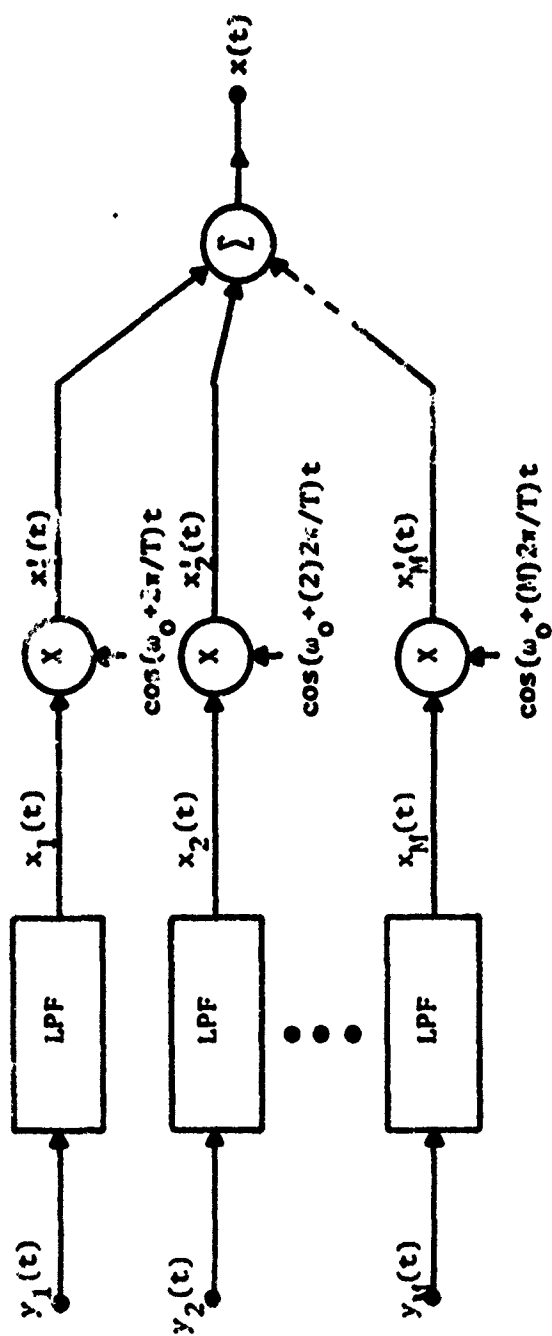


a)

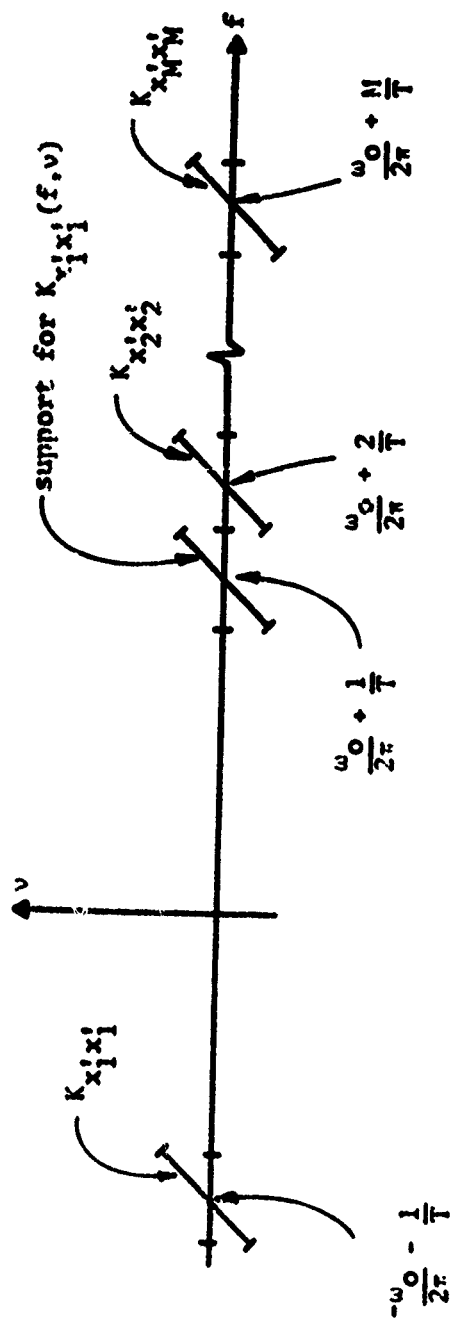


b)

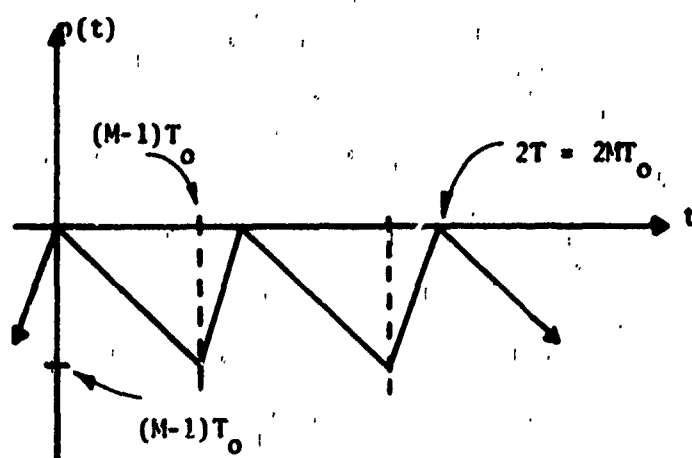
Figure(2-4) a) TDM₁ multiplexor. b) Gating signal for TDM₁.



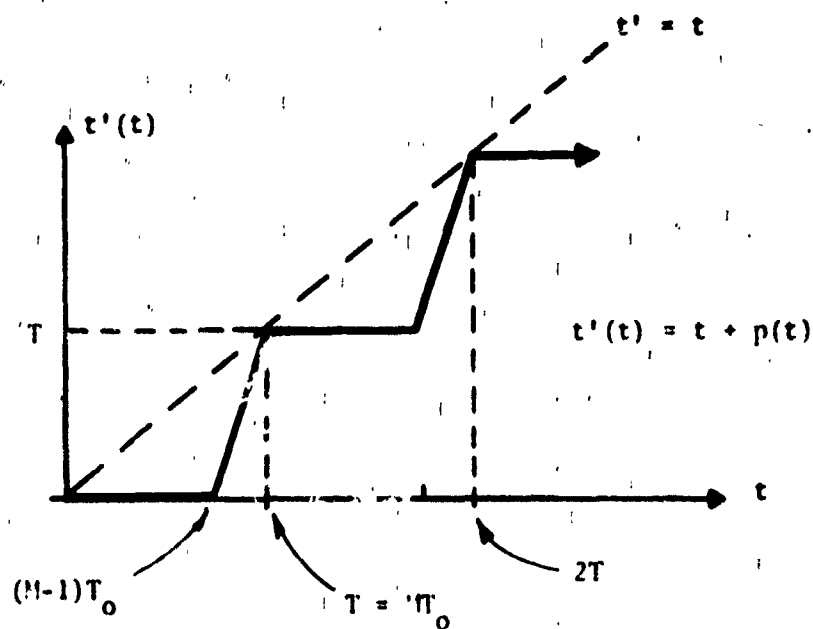
Figure(2-5a) FDM multiplexor(LPF is an ideal low-pass filter with cutoff frequency $1/2T$).



Figure(2-5b) Diagram of the support for the double Fourier transform of the autocorrelation function for an FDM signal.

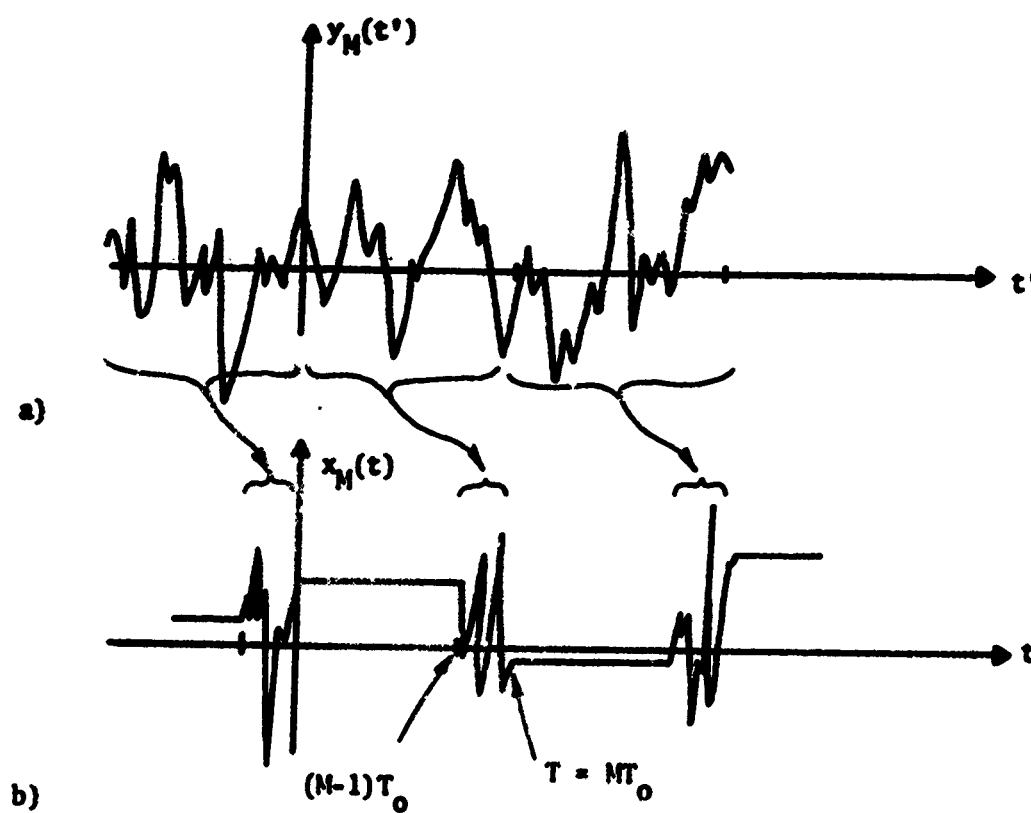


a)

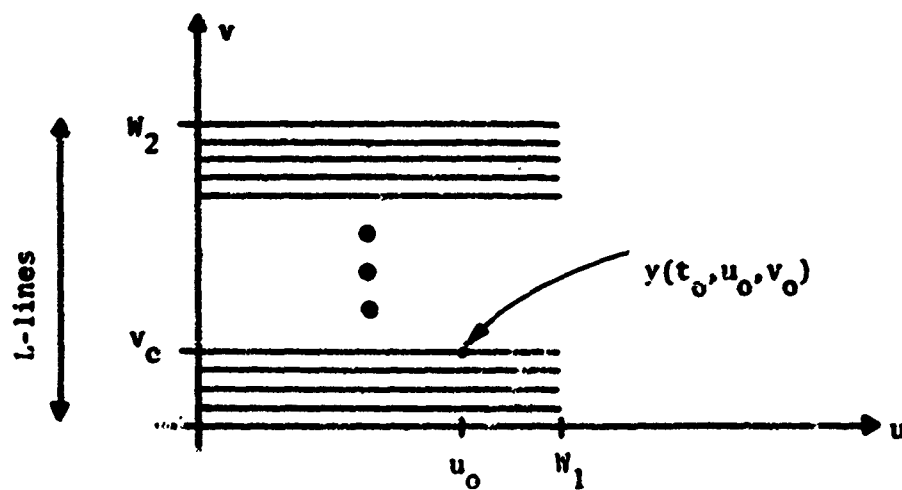


b)

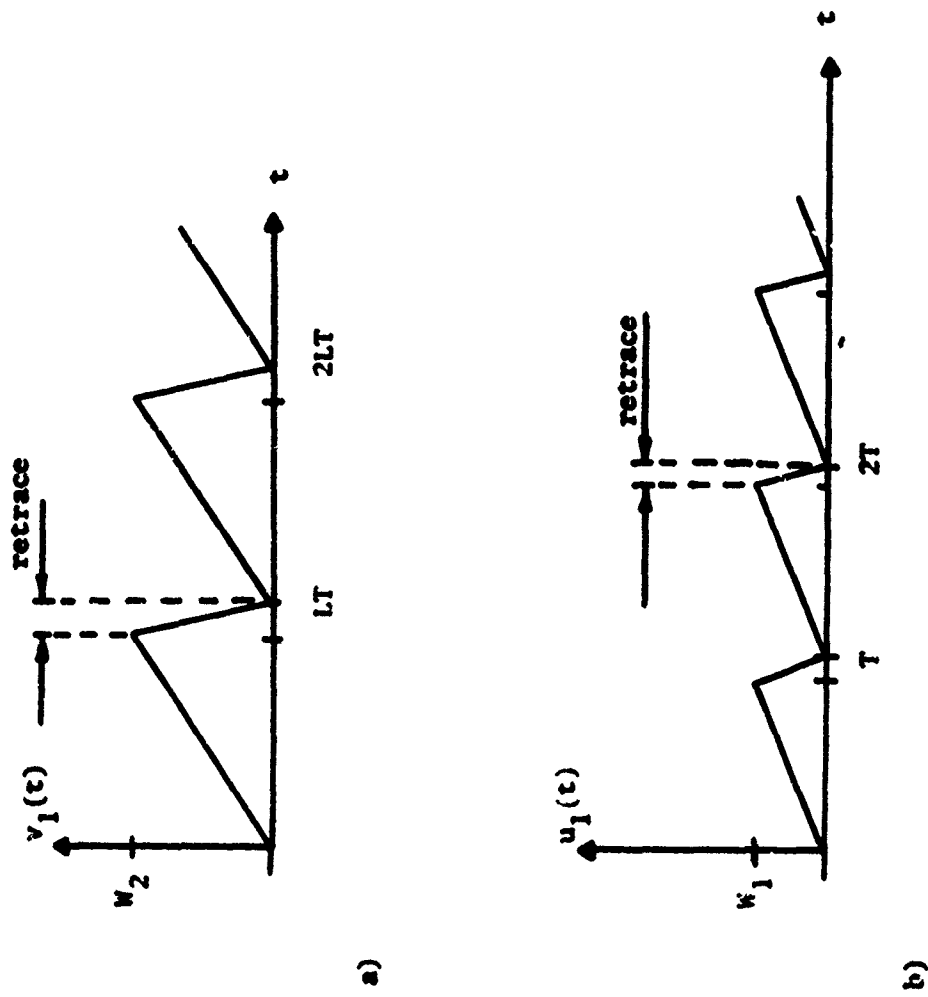
Figure(2-6) a) Periodic component of the time-scale compression transformation. b) time-scale compression transformation. (t' is compressed time, t is unaltered time.)



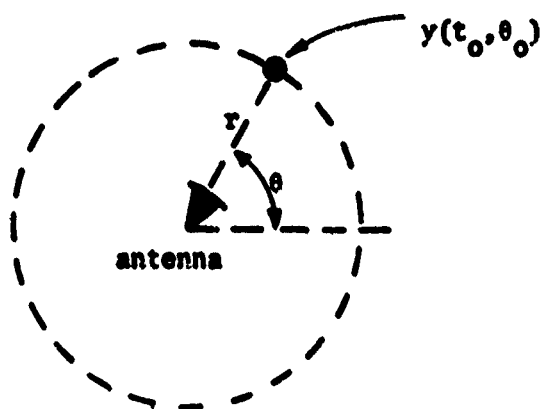
Figure(2-7) a) Unaltered process. b) Transformed process obtained by compressing T -length records into T/N -length records.



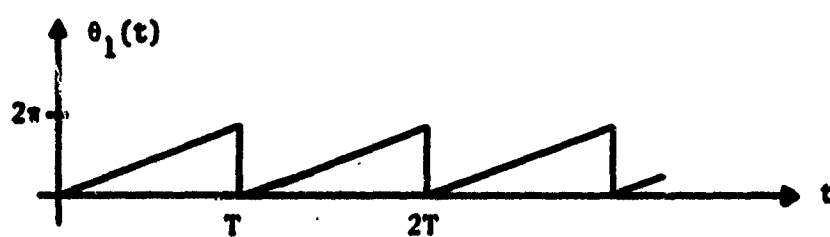
Figure(2-8) Line-scanning diagram.



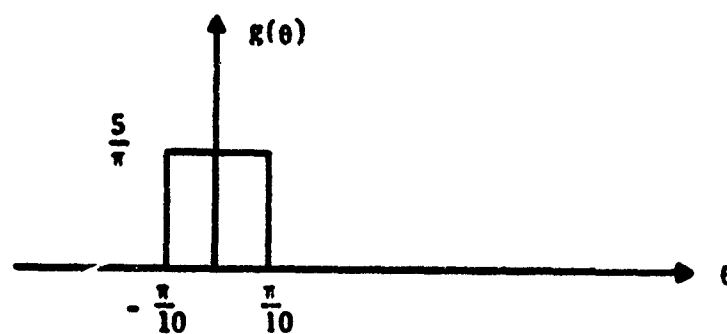
Figure(2-9) a) Vertical line-scanning function. b) Horizontal line-scanning function.



Figure(2-10) Circle-scanning diagram(r is an "effective" reception radius.)



Figure(2-11) Angular circle-scanning function.



Figure(2-12) Rectangular window function for circle-scanning.

C H A P T E R I I I

S E R I E S R E P R E S E N T A T I O N S F O R C Y C L O S T A T I O N A R Y P R O C E S S E S

(A N D T H E I R A U T O C O R R E L A T I O N F U N C T I O N S)

1. Introduction

Many analyses involving random processes are greatly facilitated by the use of appropriate representations for the processes and their autocorrelations functions. These representations are frequently categorized into one of two classes: "integral" representations and "series" representations. It is the latter class which concerns us here. In this chapter we develop two types of series representations for cyclostationary processes:

(1) Discrete "translation series" representations of the form

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t-nT) \quad \forall t \in (-\infty, \infty),$$

where $\{\phi_p\}$ are deterministic basis functions, and $\{a_{np}; p=1,2,\dots,M\}$ are jointly wide-sense-stationary discrete random processes (random sequences) referred to as "representors".

(2) Continuous series representations of the form

$$x(t) = \sum_{p=1}^M a_p(t) \tilde{\phi}_p(t) \quad \forall t \in (-\infty, \infty),$$

where $\{\tilde{\phi}_p\}$ are deterministic periodic basis functions, and $\{a_p\}$ are jointly wide-sense-stationary (or cyclostationary) continuous-time random processes referred to as the representors.

One of the advantages typically gained in representing one process in terms of a multiplicity of others is that these other processes (representors) are chosen for convenience of analysis and/or synthesis. For example, the translation series representation employs representors which are not only discrete, but also jointly wide-sense-stationary. Similarly, the continuous series representations developed in this chapter employ simple continuous-time representors such as piecewise constant processes with constant means and variances, or jointly wide-sense-stationary bandlimited processes. As shown in the next chapter, such representations enable us to convert single-variable problems with periodically varying parameters to multi-variable problems with constant parameters.

All representations presented in this chapter are valid on the infinite interval $(-\infty, \infty)$, and since all WSS processes are CS, then the representations apply to all WSS, as well as all CS, processes.

Before delving into the topic of representation, we first define a Hilbert space of random processes that will provide a framework within which we can unambiguously present our results on representation. We begin with the definitions of two well known Hilbert spaces:

(1) Hilbert space of random variables: We define H_{RV} to be the Hilbert space of finite mean-square random variables with the following inner product

$$(x, y)_{H_{RV}} \triangleq E\{xy^*\}$$

and induced norm

$$\|x\|_{H_{RV}} = [E\{|x|^2\}]^{1/2}.$$

(2) Hilbert space of functions: We define $L^2[a,b]$ to be the Hilbert space of square-integrable functions on the interval $[a,b]$ with the following inner product

$$(x,y)_{L^2} \triangleq \int_a^b x(t)y^*(t)dt$$

and induced norm

$$\|x\|_{L^2} \triangleq \left[\int_a^b |x(t)|^2 dt \right]^{1/2}.$$

We remind the reader that the elements in infinite-dimensional Hilbert spaces such as these are equivalence classes, each of which may be comprised of an uncountable number of equivalent elements, where two elements are equivalent if and only if they are equal modulo the set of zero-norm elements [2]. For example, the uncountable set

$$x_i(t) = \begin{cases} x(t), & t \neq c_i, t \in [a,b] \\ 0, & t = c_i \end{cases} \quad \forall c_i \in [a,b]$$

is an equivalence class (and therefore a single element) in $L^2[a,b]$ since the norm of the difference between any two members is zero. The terminology

$$x_i(t) = x_j(t) \quad \text{a.e. } t \in [a,b]$$

which is read " x_i equals x_j for almost every t contained in the interval $[a,b]$ " is often used to denote equivalence in $L^2[a,b]$, and will be used here occasionally; however, for the most part, the a.e. will be suppressed. In H_{RV} the terminology "mean-square equivalent" is frequently used to denote the nature of equality in that space, and will be used here occasionally.

Now, although periodic functions are not contained in the Hilbert space $L^2(-\infty, \infty)$, they can be made into a related Hilbert space of their own:

(3) Hilbert space of periodic functions: We define $L^2_{p(T)}(-\infty, \infty)$ to be the Hilbert space of square-integrable (on $[0, T]$) periodic functions on $(-\infty, \infty)$ with the commensurable periods $\{T_n; T_n = T/n, n = 1, 2, \dots\}$ with the following inner product

$$(x, y)_{L^2_p} \triangleq \int_0^T x(t) y^*(t) dt$$

and induced norm

$$\|x\|_{L^2_p} \triangleq \left[\int_0^T |x(t)|^2 dt \right]^{1/2}.$$

The cyclic character of the functions in this space is directly responsible for the existence of an L^2 -type of inner product and norm which are valid for functions on $(-\infty, \infty)$ which are not in $L^2(-\infty, \infty)$. Although this Hilbert space is a trivial variation on (and, in fact, isomorphic to) $L^2[0, T]$, it leads directly to the following--far more interesting--Hilbert space of cyclostationary processes:

(4) Hilbert space of cyclostationary processes: We define $H_{CS(T)}(-\infty, \infty)$ to be the Hilbert space of all finite mean-square jointly CS processes on $(-\infty, \infty)$ with the commensurable periods $\{T_n; T_n = T/n, n = 1, 2, \dots\}$ with the following inner product

$$(x, y)_{H_{CS}} \triangleq \int_0^T E\{x(t) y^*(t)\} dt$$

and induced norm

$$\|x\|_{H_{CS}} \triangleq \left[\int_0^T E\{|x(t)|^2\} dt \right]^{1/2}.$$

Note that, paralleling the case for $L^2_P(T)$, the existence of an L^2 -type of norm for $H_{CS}(T)$ is a direct result of the cyclic character of the second order moments of the elements in our space; i.e.,

$$E\{x(t+T)y^*(\tau+T)\} = E\{x(t)y^*(\tau)\} \quad \text{a.e. } t, \tau \in (-\infty, \infty)$$

for every x, y in $H_{CS}(T)$. When visualizing the elements of this space, recall that statistically independent and/or uncorrelated CS processes are jointly CS, and that all WSS processes are CS with arbitrary period T . In fact, the elements of $H_{CS}(T)$ might be more amenable to visualization if thought of as functions of two variables $x(t, \omega)$ where the section function $x(t, \cdot)$ is a random variable with domain: the set of events (sample space) on which the underlying probability space is defined, and with range: the complex numbers, and where the section function $x(\cdot, \omega)$ is a function in the usual sense with domain: the real numbers, and range: the complex numbers.

In concluding this introductory section, we emphasize that throughout this chapter, expressions of the form

$$x(t) = x'(t) \quad \forall t \in (-\infty, \infty)$$

--where, for example, x' is a representation for the CS process x -- should be interpreted as meaning that x and x' are in the same equivalence class in H_{CS} : i.e., the norm of their difference is zero.

2. Translation Series Representations

Due to the length of this section, and the fact that it is the heart of the theoretical content of this dissertation, we begin by giving a brief outline of the eight subsections included herein:

- Subsection: a) Definition of, and existence and identification theorem for, translation series representations.
- b) Interpretation of translation series representations as representations in terms of PAM processes.
- c) Implementation of the process-resolution and reconstruction operations.
- d) Examples of translation series representations.
- e) Classes of processes which admit finite order translation series representations.
- f) Application of translation series representations to the solution of linear integral equations.
- g) Application of translation series representations to the realization of periodically time-varying linear systems.
- h) Application of translation series representations to the definition of a generalized Fourier transform for cyclostationary (and WSS) processes.

a) Definition. A particularly interesting class of discrete representations for cyclostationary processes is the class of translation series representations.¹⁹ A cyclostationary process x with period T is said to admit a translation series representation of order M if there exists a set of M deterministic basis functions $\{\phi_p; p = 1, 2, \dots, M\}$, and a set of M jointly wide-sense-stationary random sequences $\{a_{np}; p = 1, 2, \dots, M\}$ such that

$$E\{|x(t) - \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t-nT)|^2\} = 0 \quad \text{a.e. } t \in (-\infty, \infty). \quad (3-1)$$

where, by jointly WSS, we mean that $E\{a_{np}\}$, $E\{a_{(n+m)p} a_{nq}^*\}$ are independent of n for all p, q . Thus a translation series representation (TSR) of order M is a representation of a scalar cyclostationary continuous-time process in terms of an M -vector of jointly wide-sense-stationary discrete-time processes.

The autocorrelation function for a process with a TSR of order M has the corresponding representation:

$$k_{xx}(t, s) = \sum_{n, m=-\infty}^{\infty} \sum_{p, q=1}^M A_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT) \quad \text{a.e. } t, s \in (-\infty, \infty) \quad (3-2)$$

where the elements of the $M \times M$ matrix $\{A_{n-m}\}$ of correlation sequences are given by

$$A_{n-m}^{pq} = E\{a_{np} a_{mq}^*\}, \quad (3-3)$$

¹⁹ Representations of this type have been alluded to by Jordan [25] and Breilsford [19] and perhaps some others; but have not (to my knowledge) received more than a brief mention.

and the double Fourier transform of the autocorrelation function has the representation:

$$K_{xx}(f, \nu) = \frac{1}{T} \sum_{p,q=1}^M A_{pq}(\nu) \phi_p(f) \phi_q^*(\nu) \sum_{n=-\infty}^{\infty} \delta(f - \nu + n/T), \quad (3-4)$$

where the elements of the matrix $\underline{A}(f)$ are given by

$$A_{pq}(f) = \sum_{r=-\infty}^{\infty} A_r^{pq} e^{-j2\pi r f} \quad (3-5)$$

and are z transforms of the elements (correlation sequences) of $\{\underline{A}_r\}$ evaluated at $z = e^{-j2\pi f}$, and where ϕ_p is the Fourier transform of ϕ_p .

From Eq. (3-3), we see that the matrix of correlation sequences exhibits the symmetry

$$\{\underline{A}_r\}' = \{\underline{A}_{-r}\}^*$$

so that, from Eq. (3-4), its z-transform is Hermitian symmetric:

$$\underline{A}'(f) = \underline{A}^*(f). \quad (3-6)$$

Using Theorem (2-16) of Chapter II (on the equivalence between time-averaging and phase randomizing), and Eq. (3-2), it is easily shown that the power spectral density for the stationarized version of a CS process with the above TSR has the following quadratic representation:

$$K_{xx}(f) = \frac{1}{T} \sum_{p,q=1}^M A_{pq}(f) \phi_p(f) \phi_q^*(f). \quad (3-7)$$

The type of separability exhibited by the TSR for the autocorrelation function is useful in solving integral equations--as shown in Section 2f of this chapter--which arise in estimation and detection problems (where the autocorrelation function is the kernel of a linear integral

transformation). and will be exploited in the following chapter on least-mean-square estimation.

Most CS processes of theoretical and practical interest (those contained in $H_{CS(T)}(-\infty, \infty)$) admit translation series representations of infinite order. That is, the norm of the difference between such a process and an M^{th} order approximate representation of it monotonically approaches zero as M approaches infinity. Hence, most CS processes of interest can be arbitrarily closely approximated, in norm, by a finite order TSR.

The most generally applicable type of TSR is that which employs a complete orthonormal (CON) set of basis functions $\{\phi_p\}$ in the Hilbert space $L^2[0, T]$. Orthonormality insures that, by definition,

$$(\phi_p, \phi_q)_{L^2} \triangleq \int_0^T \phi_p(t) \phi_q^*(t) dt = \delta_{pq} \triangleq \begin{cases} 1 & p = q \\ 0 & p \neq q, \end{cases} \quad (3-8)$$

and completeness insures that Bessel's equality [2] holds:

$$f(t) = \sum_{p=1}^{\infty} (f, \phi_p)_{L^2} \phi_p(t) \begin{cases} \forall f \in L^2[0, T] \\ \text{a.e. } t \in [0, T] \end{cases} \quad (3-9)$$

Note that if we extend the $\{\phi_p\}$ from $L^2[0, T]$ to $L^2(-\infty, \infty)$ by defining

$$\phi_p(t) \triangleq 0 \quad \forall t \notin [0, T],$$

then the set of translates $\{\phi_p(t-nT); p = 1, 2, \dots; n = 0, \pm 1, \pm 2, \dots\}$ is an orthonormal set in $L^2(-\infty, \infty)$ in the sense that

$$\int_{-\infty}^{\infty} \phi_p(t-nT) \phi_q^*(t-mT) dt = \delta_{pq} \delta_{nm}. \quad (3-10)$$

For this class of TSR's we have the following fundamental theorem on existence and identification:

THEOREM(3-1):²⁰ Every random process $x \in H_{CS(T)}(-\infty, \infty)$ (every finite mean-square T-CS process) admits the mean-square equivalent TSR:

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^{\infty} a_{np} \phi_p(t-nT) \quad \text{a.e. } t \in (-\infty, \infty) \quad (3-11)$$

where the elements of the jointly WSS sequences $\{a_{np}\}$ are:

$$a_{np} = \int_0^T x(t+nT) \phi_p^*(t) dt \quad (3-12)$$

and where $\{\phi_p\}$ is any CON set in $L^2[0, T]$.

Proof: Define

$$e(t) \triangleq E\{|x(t) - \sum_{n=-\infty}^{\infty} \sum_{p=1}^{\infty} a_{np} \phi_p(t-nT)|^2\}.$$

Expanding the square and interchanging summation and expectation yields:

$$\begin{aligned} e(t) = k_{xx}(t, t) - 2\text{Re}\left[\sum_n \sum_p E\{a_{np} x^*(t)\} \phi_p(t-nT)\right] \\ + \sum_{n,m} \sum_{p,q} E\{a_{np} a_{mq}^*\} \phi_p(t-nT) \phi_q^*(t-mT). \end{aligned}$$

Substituting the given expression for a_{np} and interchanging integration and expectation yields:

²⁰ Note that any generally nonstationary finite mean-square random process can be represented on $(-\infty, \infty)$ with a TSR. The discrete representors will not, in general, be jointly WSS, but the representation will be a genuine discrete series representation in terms of a countable number of deterministic basis functions with random-variable weights. However, in contrast to the usual discrete representations on a finite interval, the TSR on $(-\infty, \infty)$ cannot be truncated at a finite number of terms to provide an arbitrarily close (in mean-square) approximation (because the sum over the translation index must always be from $-\infty$ to ∞ if the representation is to apply over the entire interval $(-\infty, \infty)$).

$$\begin{aligned}
e(t) = & k_{xx}(t, t) - 2\operatorname{Re} \left[\sum_n \sum_p \int_0^T k_{xx}(\sigma+nT, t) \phi_p^*(\sigma) d\sigma \phi_p(t-nT) \right] \\
& + \sum_{n,m} \sum_{p,q} \int_0^T \int_0^T k_{xx}(\sigma+nT, \tau+mT) \phi_p^*(\sigma) \phi_q(\tau) d\sigma d\tau \phi_p(t-nT) \phi_q^*(t-mT)
\end{aligned}$$

or, changing the variables of integration:

$$\begin{aligned}
e(t) = & k_{xx}(t, t) - 2\operatorname{Re} \left[\sum_n \sum_p \int_{nT}^{(n+1)T} k_{xx}(\sigma, t) \phi_p^*(\sigma-nT) d\sigma \phi_p(t-nT) \right] \\
& + \sum_{n,m} \sum_{p,q} \int_{nT}^{(n+1)T} \int_{mT}^{(m+1)T} k_{xx}(\sigma, \tau) \phi_p^*(\sigma-nT) \phi_q(\tau-mT) d\sigma d\tau \cdot \\
& \phi_p(t-nT) \phi_q^*(t-mT). \quad (3-13)
\end{aligned}$$

Now, since x is a finite mean-square process, then the section function $k_{xx}(\cdot, t)$ is in $L^2[0, T]$ for every $t \in (-\infty, \infty)$, and since x is T -CS, then k_{xx} is jointly T -periodic in its two arguments so that $k_{xx}(\cdot, t)$ is in $L^2[nT, (n+1)T]$ for all integers n and every $t \in (-\infty, \infty)$. Also, by hypothesis $\{\phi_p(t-nT)\}$ are CON in $L^2[nT, (n+1)T]$. Hence, from Bessel's equality (Eq. (3-9)), we have:

$$\sum_{p=1}^{\infty} \int_{nT}^{(n+1)T} k_{xx}(\sigma, t) \phi_p^*(\sigma-nT) d\sigma \phi_p(t'-nT) = k_{xx}(t', t)$$

for all n , for a.e. $t' \in [nT, (n+1)T]$, and for a.e. $t \in (-\infty, \infty)$, and the left member is identically zero for $t' \notin [nT, (n+1)T]$. Hence,

$$\sum_{n=-\infty}^{\infty} \sum_{p=1}^M \int_{nT}^{(n+1)T} k_{xx}(\sigma, t) \phi_p^*(\sigma-nT) d\sigma \phi_p(t'-nT) = k_{xx}(t', t) \quad (3-14)$$

for a.e. $t, t' \in (-\infty, \infty)$.

Similarly,

$$\begin{aligned}
& \sum_m \sum_q \int_{mT}^{(m+1)T} \left[\sum_n \sum_p \int_{nT}^{(n+1)T} k_{xx}(\sigma, \tau) \phi_p^*(\sigma-nT) d\sigma \phi_p(t'-nT) \right] \cdot \\
& \phi_q(\tau-mT) d\tau \phi_q^*(t-mT) \\
& = \sum_m \sum_q \int_{mT}^{(m+1)T} k_{xx}(t', \tau) \phi_q(\tau-mT) d\tau \phi_q^*(t-mT)
\end{aligned}$$

$$= k_{xx}(t', t) \quad (3-15)$$

for a.e. $t, t' \in (-\infty, \infty)$.

Now, substituting Eqs. (3-14,15) into Eq. (3-13) yields:

$$e(t) = k_{xx}(t, t) - 2k_{xx}(t, t) + k_{xx}(t, t) = 0 \quad \text{a.e. } t \in (-\infty, \infty).$$

Hence, x and its TSR are mean-square equivalent for a.e. $t \in (-\infty, \infty)$.

The fact that the sequences $\{a_{np}\}$ are jointly WSS can be trivially demonstrated using their defining equation (Eq. (3-12)) and the fact that x is T-CS.

QED

Note that if we substitute Eq. (3-12) of this theorem into Eq. (3-3), then we obtain the following direct formulas for the correlation matrices of the jointly WSS sequences:

$$\Lambda_r^{pq} = \int_0^T \int k_{xx}(t+rT, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau \quad (3-16)$$

and, using Parseval's relation [2]:

$$\Lambda_r^{pq} = \int_{-\infty}^{\infty} \int k_{xx}(f, \nu) \phi_p^*(f) \phi_q(\nu) e^{j2\pi r T f} df d\nu. \quad (3-17)$$

b) Representation in terms of PAM processes. An interesting interpretation of the TSR of Theorem (3-1) results from the following manipulation:

Define the periodic extension of ϕ_p as

$$\tilde{\phi}_p(t) \triangleq \sum_{n=-\infty}^{\infty} \phi_p(t-nT) \quad (3-18)$$

and define a gate function w as

$$w(t) = \begin{cases} 1, & t \in [0, T] \\ 0, & t \notin [0, T] \end{cases} \quad (3-19)$$

Now, we have

$$\phi_p(t) = \tilde{\phi}_p(t)w(t),$$

and we can write the discrete M^{th} order TSR for x as an M -term continuous series representation as follows:

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t-nT) \\ &= \sum_p \sum_n a_{np} w(t-nT) \tilde{\phi}_p(t) \\ &= \sum_{p=1}^M a_p(t) \tilde{\phi}_p(t), \end{aligned} \quad (3-20)$$

where $\{a_p(t)\}$ are pulse-amplitude modulated (PAM) processes with full duty-cycle rectangular pulses:

$$a_p(t) = \sum_{n=-\infty}^{\infty} a_{np} w(t-nT). \quad (3-21)$$

A typical realization of one of these PAM processes is shown in Fig. (1-2) of Chapter I.

The autocorrelation function for x has the corresponding representation

$$k_{xx}(t, s) = \sum_{p, q=1}^M A_{pq}(t, s) \tilde{\phi}_p(t) \tilde{\phi}_q^*(s), \quad (3-22)$$

where

$$\begin{aligned} A_{pq}(t, s) &= E\{a_p(t) a_q^*(s)\} \\ &= \sum_{n, m} A_{n-m}^{pq} w(t-nT) w(s-mT), \end{aligned} \quad (3-23)$$

and since the correlation matrix $\underline{A}(t,s)$ is jointly T-periodic in its two arguments, it is easily shown that the M PAM processes $\{a_p(t)\}$ are jointly T-CS. Furthermore, as pointed out in Chapter II, these PAM representors have constant means and variances.

This representation would be even more interesting if the PAM processes $\{a_p(t)\}$ turned out to be jointly WSS. Using Eq. (3-23) it can be shown that this will, in fact, be the case if and only if the Fourier transform of w --rather than w itself--is a gate function (on the interval $[-1/2T, 1/2T]$). With such a choice of w , we obtain the "harmonic series" representation which is defined and discussed at length in Section 3.

c) Implementation of the process-resolution operation. Knowledge of Eq. (3-12) in Theorem (3-1) allows us to explicitly specify and, in fact, implement the process-resolution operation; i.e., the operation which resolves the random process x into its representors $\{a_{np}\}$. The operation indicated by Eq. (3-12) has two realizations--both of which are made causal [2] by the incorporation of a delay of length T: the multiplier-integrator realization of Fig. (3-1), and the filter-sampler realization of Fig. (3-2).

The multiplier-integrator resolution device, as shown in Fig. (3-1), is useful only for computing a single representor random variable, since the integrator must be "dumped", after computation of a_{np} , before it can be used to compute $a_{(n+1)p}$. But this can be easily remedied since an effect which is equivalent to dumping can be obtained by following the sampled output of the integrator with a feedback-delay as shown in Fig. (3-3). Furthermore, if the sampler at the output of the integrator is replaced with a sample-and-hold device (where the sample value

obtained at time nT is held until time $(n+1)T$ so that the sampler's output is piecewise constant), as shown in Fig. (3-4), then the output of the feedback-delay device is the continuous representor (PAM process) $a_p(t)$, and its sequence of sample values is the discrete representor $\{a_{np}\} = \{a_p(nT)\}$.

Now, using M of these resolution devices, we can construct a causal, linear, periodically time-varying system, as shown in Fig. (3-5), which resolves the process x into its representors. Furthermore, we can reconstruct x from these representors with the causal, linear, periodically-time-varying system also shown in Fig. (3-5).

In the event that the $\{\phi_p\}$ are complex and indexed from $-(M-1)/2$ to $(M-1)/2$ such that $\phi_p^* = \phi_{-p}$, then the resolution-reconstruction system of Fig. (3-5) can be equivalently realized using only real devices as shown in Fig. (3-6). This system computes, separately, the real and imaginary parts of the representors.

Unlike the multiplier-integrator device of Fig. (3-1), the filter-sampler device of Fig. (3-2) can be used without modification to continuously compute representors, so that the composite resolution-reconstruction system takes the form shown in Fig. (3-7).

d) Examples.

i) Harmonic translation-series-representations. If we choose as our CON set in $L^2[0,T]$ the complex exponentials

$$\phi_p(t) = \frac{1}{\sqrt{T}} e^{j2\pi pt/T} w(t)$$

where w is the gate function defined in Eq. (3-19), then the periodic extension of ϕ_p is just

$$\phi_p(t) = \frac{1}{\sqrt{T}} e^{j2\pi pt/T}$$

and the PAM representation of Eq. (3-19) becomes:

$$x(t) = \frac{1}{\sqrt{T}} \sum_{p=(M-1)/2}^{(M+1)/2} a_p(t) e^{j2\pi pt/T}$$

where the p^{th} PAM process a_p is given by

$$a_p(t) = \sum_n a_{np} w(t-nT)$$

and has pulse-amplitudes given by

$$a_{np} = \frac{1}{\sqrt{T}} \int_0^T x(t+nT) e^{-j2\pi pt/T} dt.$$

For this harmonic representation, the multiplier function employed in the resolution-reconstruction system shown in Fig. (3-6) are simply sinusoids given by

$$\operatorname{Re}\{\phi_p(t)\} = \frac{1}{\sqrt{T}} \cos(2\pi pt/T)$$

$$\operatorname{Im}\{\phi_p(t)\} = \frac{1}{\sqrt{T}} \sin(2\pi pt/T).$$

Note that if $M = 1$, then the process x is itself a PAM process with full duty-cycle rectangular pulses and pulse-amplitudes $\{x(nT)\}$.

ii) Walsh translation-series-representations. An interesting choice--from the point of view of implementation--for our CON set in $L^2[0,T]$ is the set of two-level Walsh functions [2] defined as

$$\phi_p(t) = \begin{cases} 1/\sqrt{T}, & \forall t \in I_p \\ -1/\sqrt{T}, & \forall t \in [0,T] - I_p \end{cases}$$

where I_p is the set of disjoint intervals contained in $[0,T]$ over which ϕ_p is positive. The first few sets are

$$I_1 = [0, T]$$

$$I_2 = [0, T/2]$$

$$I_3 = [0, T/4], [3T/4, T]$$

$$I_4 = [0, T/4], [T/2, 3T/4].$$

Now, since multiplication of a process by a Walsh function is equivalent to amplitude scaling by the factor $1/\sqrt{T}$ followed by a series of polarity reversals, then the multipliers shown in the representation system of Fig. (3-5) can be implemented as polarity reversal switches which are actuated at the transition times of the Walsh functions (end points of the intervals in the sets I_p). Hence, the overall implementation of the resolution-reconstruction system can be very simple.

Note that this Walsh representation (Eq. (3-20)) is composed of the sum of M piecewise constant processes, and is therefore also piecewise constant (except in the limit when $M = \infty$).

iii) Karhunen-loève translation-series-representation. The basis functions chosen for our first two examples--sinusoids and two-level functions--are appealing because of the resultant convenience in implementation of the multiplier function generators and/or the multiplier-devices used in the resolution and reconstruction systems. However, there is another criterion which brings to our attention a third choice for basis functions. Consider the choice of a set of basis functions which, for a given process x and any given order of approximate representation M , minimizes the integral-mean-squared difference between the process x and its approximate representation x' . In this case, our choice would be appealing from the point of view of efficiency.

Now, it is well known [2,25] that the set of orthonormalized eigenfunctions $\{\phi_p; p = 1, 2, \dots, M\}$, corresponding to the M largest eigenvalues $\{\lambda_p\}$ of the linear integral operator on $L^2[0, T]$ with kernel $k_{xx}(t, s)$, is the unique set which minimizes the integral-mean-squared difference:

$$\int_0^T E\{[x(t) - \sum_{p=1}^M a_{0p} \phi_p(t)]^2\} dt, \quad a_{0p} = (x, \phi_p)_{L^2}.$$

The series representation which employs these eigenfunctions

$$x'(t) = \sum_{p=1}^M a_{0p} \phi_p(t), \quad t \in [0, T]$$

is referred to as the Karhunen-Loève representation for x on $[0, T]$ and is the optimum (most efficient) discrete representation. The eigenfunctions are, by definition, the solutions to the integral equation:

$$\lambda_p \phi_p(t) = \int_0^T k_{xx}(t, s) \phi_p(s) ds, \quad \forall t \in [0, T], \quad (3-25)$$

and if x is T -CS, then they are also the solutions to the integral equation:

$$\lambda_p \phi_p(t - nT) = \int_{nT}^{(n+1)T} k_{xx}(t, s) \phi_p(s - nT) ds, \quad \forall t \in [nT, (n+1)T] \quad (3-26)$$

Hence, the series representation

$$x'(t) = \sum_{p=1}^M a_{np} \phi_p(t - nT), \quad \forall t \in [nT, (n+1)T]$$

is also the Karhunen-Loève representation for x on $[nT, (n+1)T]$ so that the resultant TSR

$$x'(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t - nT), \quad \forall t \in (-\infty, \infty)$$

minimizes the integral-mean-squared difference in every interval $[nT, (n+1)T]$, and is therefore the most efficient TSR for x .

For this Karhunen-Loève TSR, the matrix \underline{A}_0 with elements $\{A_0^{pq}\}$ is, as expected, diagonal; but the matrices $\{\underline{A}_r; r \neq 0\}$ are not, in general diagonal. However, there exists an interesting class of autocorrelation functions for which the latter matrices are diagonal and, in fact, scalar multiples of \underline{A}_0 . For details, see Theorem (3-3) and the following discussion.

We now present explicit formulas for the eigenfunctions and eigenvalues employed in the Karhunen-Loève TSR for two types of CS processes:

(1) Karhunen-Loève TSR for the video signal: In Section 3 of Chapter II, we presented a model for the random video signal which included a formula for the autocorrelation function consisting of three factors:

$$k_{xx}(t,s) = k_1(t-s)k_2(t,s)k_3(t,s).$$

If we ignore the frame-to-frame correlation (because of its relatively low-frequency nature) by letting ρ_3 in Eq. (2-26) equal 1 (and $L \rightarrow \infty$), then $k_3(t,s) = 1$ for all t,s and the video signal becomes CS with period T rather than LT . Furthermore, on the square $t,s \in [0,T]$, the line-to-line correlation factor $k_2(t,s) = 1$ (from Eq. (2-25)). Thus, we have

$$k_{xx}(t,s) = k_1(t-s), \quad \forall t,s \in [0,T]$$

so that, from Eq. (2-24),

$$k_{xx}(t,s) = \rho_1 \frac{|t-s|}{T}, \quad \forall t,s \in [0,T].$$

Now, from Eq. (3-25), our TSR basis functions are the normalized solutions of the eigenfunction equation:

$$\lambda_p \phi_p(t) = \int_0^T \rho_1 |t-s| \phi_p(s) ds, \quad \forall t \in [0, T],$$

and are given (unnormalized) by [2]

$$\phi_p(t) = \begin{cases} \cos(\pi f_0 (t-T/2) \gamma_p) & , \quad p \text{ odd} \\ \sin(\pi f_0 (t-T/2) \gamma_p) & , \quad p \text{ even} \end{cases}$$

and

$$\lambda_p = \frac{1/\pi f_0}{1 + \gamma_p^2}, \quad (3-27)$$

where $\{\gamma_p\}$ are the solutions of

$$\tan(\pi f_0 T \gamma_p) = \begin{cases} 1/\gamma_p, & p \text{ odd} \\ -\gamma_p, & p \text{ even} \end{cases}$$

and $f_0 \triangleq \frac{1}{2\pi} \ln(1/\rho_1)$.

(2) Karhunen-Loève TSR for the time-division-multiplexed signal:

In Section 2 of Chapter II, we presented a model for the time-division-multiplex (TDM₁) of M random signals $\{y_p\}$. If we assume that these M component signals are uncorrelated and WSS with autocorrelation functions $\{k_p\}$, then the composite autocorrelation function for the TDM₁ signal x becomes:

$$k_{xx}(t, s) = \sum_{n, m=-\infty}^{\infty} \sum_{p=1}^M k_p(t-s) w_p(t-nT) w_p(s-mT) \quad (3-28)$$

where w_p is the gate function

$$w_p(t) = \begin{cases} 1, & \forall t \in [(p-1)T/M, pT/M] \triangleq T_p \\ 0, & \forall t \notin T_p \end{cases},$$

so that, on the square $t, s \in [0, T]$, we have

$$k_{xx}(t, s) = \sum_{p=1}^M k_p(t-s) w_p(t) w_p(s), \quad \forall t, s \in [0, T],$$

and on the square $t, s \in T_p$, we have

$$k_{xx}(t, s) = k_p(t-s), \quad \forall t, s \in T_p.$$

Now, the most convenient way to obtain a TSR for x is to represent x separately on each of the M component intervals $\{T_p\}$ with an N^{th} order Karhunen-Loève representation:

$$x'(t) = \sum_{q=1}^N a_q^P \phi_q^P(t) \quad \forall t \in T_p.$$

This will be equivalent to representing each of the M component processes $\{y_p\}$ individually:

$$y_p'(t) = \sum_{q=1}^N a_q^P \phi_q^P(t), \quad \forall t \in T_p.$$

The result is the composite $(M \cdot N)^{\text{th}}$ order TSR:

$$x'(t) = \sum_{p=1}^M \sum_{q=1}^N \sum_{n=-\infty}^{\infty} a_{nq}^P \phi_q^P(t-nT), \quad \forall t \in (-\infty, \infty). \quad (3-29)$$

If the component processes $\{y_p\}$ have exponential autocorrelation functions

$$k_p(t-s) = e^{-2\pi f_p |t-s|},$$

then the eigenfunctions $\{\phi_q^P\}$ can be obtained from the formula given in the Video example. Another interesting type of autocorrelation function which, unlike the exponential type, is duration limited, is that which takes the form

$$k(t-s) = \begin{cases} \frac{k(0)}{T_*^n} [(t-s+T_*)^n u(s-t) + (s-t+T_*)^n u(t-s)], & |t-s| \leq T_*, \text{ } n \text{ odd} \\ 0, & |t-s| \geq T_* \end{cases}$$

and has the shape shown in Fig. (3-8).

The eigenfunctions on the interval $[-T_0/2, T_0/2]$ (centered for convenience) which correspond to this type of kernel are very easy to solve for provided that

$$T_0 \leq T_*/2.$$

The approach, like that used to obtain the eigenfunctions for exponential kernels [2], relies on converting the integral equation Eq. (3-25) to a differential equation. For the above type of kernel, the differential equation is:

$$\lambda_p \left[\frac{d}{dt} \right]^{n+1} \phi_p(t) = -\frac{2k(0)n!}{T_*^n} \phi_p(t), \quad \forall t \in [-T_0/2, T_0/2], \text{ } n \text{ odd.}$$

For the case $n = 1$, the kernel (autocorrelation function) is triangular, and the eigenfunctions are given by:

$$\phi_p(t) = \cos(\omega_p t),$$

and the eigenvalues by

$$\lambda_p = 2k(0)/T_* \omega_p^2, \quad (3-30)$$

where the $\{\omega_p\}$ are the solutions of

$$\tan(\omega_p T_0) = \frac{1}{\omega_p T_0} - \frac{T_*}{T_0} [1/(\omega_p T_0)^2 + 1]^{1/2}.$$

e) Finite order translation series representations. The TSR's established in Theorem (3-1) are generally applicable to all CS processes with bounded autocorrelation functions, but will usually be of infinite order ($M = \infty$). There are, on the other hand, a number of subclasses of CS processes which admit special finite order TSR's which are not based on some CON set in $L^2[0,T]$ as in Theorem (3-1). We present two such subclasses.

i) Bandlimited cyclostationary processes. In many analyses dealing with random signals (particularly in communications studies) all signals are considered to be bandlimited to the interval of frequencies $[-f_0, f_0]$ where f_0 is the highest frequency which the systems (in the problem of interest) will respond to. Thus, given f_0 and a T-CS process x , we choose the smallest integer M such that $M/2T \geq f_0$, and then remodel our process as its bandlimited (to $[-M/2T, M/2T]$) version:

$$x'(t) = \int_{-\infty}^{\infty} x(\tau) \frac{\sin((t-\tau)\pi M/T)}{(t-\tau)\pi M/T} d\tau.$$

Now, from the sampling theorem for nonstationary processes (Theorem (2-2)), we have the following mean-square equivalent representation for any process x which is bandlimited to $[-M/2T, M/2T]$:

$$x(t) = \sum_{m=-\infty}^{\infty} x(mT/M) \frac{\sin((t-mT/M)\pi M/T)}{(t-mT/M)\pi M/T}.$$

Thus, if we define the M basis functions $\{\phi_p\}$ as

$$\phi_p(t) \triangleq \left(\frac{M}{T}\right)^{1/2} \frac{\sin((t-pT/M)\pi M/T)}{(t-pT/M)\pi M/T} \quad (3-31)$$

and the M representor sequences $\{a_{np}\}$ as

$$a_{np} \triangleq \left(\frac{T}{M}\right)^{1/2} x(nT + pT/M), \quad (3-32)$$

then we can rewrite the above sampling representation as an M^{th} order TSR

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t-nT) \quad (3-33)$$

where the $\{\phi_p(t-nT)\}$ are orthonormal on $L^2[-\infty, \infty]$ (as in Eq. (3-10)), and--analogous to Eq. (3-12) of Theorem (3-1)--the elements of the M jointly WSS sequences are given by

$$a_{np} = \int_{-\infty}^{\infty} x(t+nT) \phi_p(t) dt. \quad (3-34)$$

ii) Degenerate cyclostationary processes. We use the term "degenerate" to denote those random processes which have a countable (cardinality equal to that of the set of natural numbers) ensemble of realizations. We consider here a specific subclass of degenerate CS processes which are of much practical interest in communications: synchronous M -ary signals of the form

$$x(t) = \sum_{n=-\infty}^{\infty} \phi(t-nT, b_n),$$

where ϕ is a deterministic function and $\{b_n\}$ is an M -ary random sequence where each random variable b_n has the M -ary alphabet of realizations $\{\alpha_1, \alpha_2, \dots, \alpha_M\}$. In Section 4 of Chapter II, we defined this type of CS process and gave six specific examples including frequency-shift-keyed signals, pulse-position-modulated signals, and others.

Now, such processes can be re-expressed as an M^{th} order TSR as in Eq. (3-33) where the p^{th} basis function is

$$\phi_p(t) = \phi(t, \alpha_p),$$

and where the M jointly WSS sequences $\{a_{np}\}$ comprise a random indicator sequence. That is, for each n , the realizations of all but one of the M elements $\{a_{n1}, a_{n2}, \dots, a_{nM}\}$ are zero, and the non-zero realization is equal to 1. For this TSR, the matrix of correlation sequences, $A_{n-m}^{pq} \triangleq E\{a_{np} a_{mq}^*\}$, is a matrix of joint probabilities of which the n - m th element of the pq th sequence is the joint probability that $b_n = \alpha_p$ and $b_m = \alpha_q$. Furthermore, the mean value of the random variable a_{np} is just the probability that $b_n = \alpha_p$.

f) Correlation matrix decomposition and solution of integral equations.

In this subsection, we show how the TSR for an autocorrelation function can be used to obtain the solution of a linear integral equation whose kernel is the autocorrelation function, and we show how to simplify this solution through decomposition of the correlation matrix of Eq. (3-5).

i) Integral equations. The solutions to many problems in functional analysis (including estimation and detection problems as discussed in Chapter IV) can ultimately be expressed in terms of the solution to a linear integral equation--typically a Fredholm equation of the first kind [29,31]

$$\int_S k(t,s)h(s)ds = g(t), \quad \forall t \in S, \quad (3-35)$$

or of the second kind

$$\int_S k(t,s)h(s)ds + \lambda h(t) = g(t), \quad \forall t \in S, \quad (3-36)$$

or, more generally,

$$\int_S k(t,s)h(s,\tau)ds = g(t,\tau), \quad \forall t,\tau \in S, \quad (3-37)$$

or

$$\int_S k(t,s)h(s,\tau)ds + \lambda h(t,\tau) = g(t,\tau), \quad \forall t,\tau \in S, \quad (3-38)$$

where S is some fixed indexing interval.

Since the first three equations can be considered to be special cases of the fourth, we will restrict our discussion here to this fourth equation.

There are various sets of circumstances under which explicit solutions to this integral equation can be obtained. For example:

(1) If $S = (-\infty, \infty)$, and if the kernel k and right member g are time-invariant ($k(t,s) = k(t-s)$), then the integral is a convolution and the solution is (formally):

$$h(t,s) = F^{-1} \left\{ \frac{G}{K + \lambda} \right\}_{t-s}$$

where $F^{-1}(\cdot)_{t-s}$ is the inverse Fourier transform evaluated at $t-s$, and G, K are the Fourier transforms of g, k .

(2) If the kernel k and right member g are "separable of order M ":

$$\begin{aligned} k(t,s) &= \sum_{p,q=1}^M A_{pq} \phi_p(t) \phi_q(s) \\ g(t,s) &= \sum_{p,q=1}^M G_{pq} \phi_p(t) \phi_q(s) \quad \forall t,s \in S \end{aligned}$$

($\{\phi_p\}$ can always be chosen orthonormal by proper choice of the matrices A, G), then the integral transformation is degenerate, and the solution is (formally):

$$h(t,s) = \sum_{p,q=1}^M B_{pq} \phi_p(t) \phi_q(s)$$

where the matrix \underline{B} is given by

$$\underline{B} = [\underline{A} + \lambda \underline{I}]^{-1} \underline{G}$$

where \underline{I} is the $M \times M$ identity matrix.

Note, if S is finite, then under a fairly liberal set of constraints on k, g , arbitrarily close separable approximations to k, g can be obtained [34,36].

Now, as a combination of these two methods, we present the following new (to our knowledge) formal solution to Eq. (3-38):

THEOREM(3-2): If the symmetric functions k, g are each jointly T -periodic in their two arguments ($k(t+T, s+T) = k(t, s)$), and if the section functions $k(\cdot, t), k(t, \cdot), g(\cdot, t), g(t, \cdot)$ are in $L^2[0, T]$ for every $t \in (-\infty, \infty)$, then the solution to the integral equation

$$\int_{-\infty}^{\infty} k(t, s) h(s, \tau) ds + \lambda h(t, \tau) = g(t, \tau) \quad \forall t, \tau \in (-\infty, \infty) \quad (3-39)$$

has the representation:

$$h(t, \tau) = \sum_{n,m=-\infty}^{\infty} \sum_{p,q=1}^{\infty} B_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(\tau-mT) \quad (3-40)$$

where $\{\phi_p\}$ is any CON set on $L^2[0, T]$, and the sequence of matrices $\{\underline{B}_r\}$ is given by the inverse z -transform

$$\underline{B}_r = T \int_{-1/2T}^{1/2T} \underline{B}(f) e^{j2\pi r T f} df \quad (3-41)$$

where the matrix of functions $\underline{B}(f)$ is given by

$$\underline{B}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1} \underline{G}(f) \quad (3-42)$$

where the matrices \underline{A} , \underline{G} are the z-transforms

$$\begin{aligned} \underline{A}(f) &= \sum_{r=-\infty}^{\infty} \underline{A}_r e^{-j2\pi r T f} \\ \underline{G}(f) &= \sum_{r=-\infty}^{\infty} \underline{G}_r e^{-j2\pi r T f} \end{aligned} \quad (3-43)$$

and where the elements of the sequences of matrices \underline{A}_r , \underline{G}_r are given by

$$\begin{aligned} A_r^{pq} &= \int_0^T \int_0^T k(t+rT, \tau) \phi_p^*(\tau) \phi_q(\tau) d\tau dt \\ G_r^{pq} &= \int_0^T \int_0^T g(t+rT, \tau) \phi_p^*(\tau) \phi_q(\tau) d\tau dt. \end{aligned} \quad (3-44)$$

Proof: Employing the completeness relation (Eq. (3-9)) as in the proof of Theorem (3-1) (Eq. (3-15)) yields the translation series representations

$$\begin{aligned} k(t, s) &= \sum_{n, m=-\infty}^{\infty} \sum_{p, q=1}^{\infty} A_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT), \\ g(t, s) &= \sum_{n, m} \sum_{p, q} G_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT). \end{aligned} \quad (3-45)$$

Now, assuming the form Eq. (3-40) for h , and substituting these representations into Eq. (3-39), and employing the orthonormality of the $\{\phi_p(t-nT)\}$ (Eq. (3-10)) yields the discrete version of our integral equation Eq. (3-39):

$$\sum_{l=1}^{\infty} \sum_{m=-\infty}^{\infty} A_{n-m}^{pl} B_m^{lq} + \lambda B_n^{pq} = G_n^{pq} \quad \forall n, p, q. \quad (3-46)$$

But, the first term in the left member is simply the composition of a discrete convolution and a matrix product. Recognizing this and taking the z-transform of both sides of this equation

yields

$$\underline{A}(f) \cdot \underline{B}(f) + \lambda \underline{B}(f) = \underline{G}(f)$$

so that (formally)

$$\underline{B}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1} \underline{G}(f).$$

Now, the sequences $\{B_r^{pq}\}$ are the inverse z-transforms given in Eq. (3-41).

QED

Note that the matrix $\underline{B}(f)$ which characterizes the solution given in this theorem is defined in terms of the inverse matrix $[\underline{A}(f) + \lambda \underline{I}]^{-1}$ where $\underline{A}(f)$ is, in general, infinite dimensional. Thus, our solution is, in general, only formal. However, in the event that finite order (say M) translation series representations for k, g exist, then the matrices \underline{A} , \underline{B} , \underline{G} are all of dimension $M \times M$ and the solution can be of great value as illustrated in Chapter IV. Furthermore, even when $\underline{A}(f)$ is infinite dimensional, the inverse $[\underline{A}(f) + \lambda \underline{I}]^{-1}$ can be computed provided $\underline{A}(f)$ is either diagonal or can be appropriately decomposed into the sum of a diagonal matrix and a finite rank matrix. In this next subsection, we discuss the decomposition of $\underline{A}(f)$ and give a number of examples.

However, before turning to this topic, we note that the solution method of Theorem (3-2) can easily be extended (as shown in Chapter IV) to the case where the basis functions $\{\phi_p\}$ are neither orthonormal nor duration limited to $[0, T]$.

ii) Correlation matrix decomposition. In the previous section on integral equations, we saw (Theorem (3-2)) that an operator equation of the form

$$[\underline{A} - \lambda \underline{I}] \cdot \underline{B} = \underline{G}, \quad (3-47)$$

where \underline{A} , \underline{B} , \underline{G} are operators on the function space $L^2(-\infty, \infty)$ with kernels k, h, g which admit a TSR of order M (Eq. (3-45)):

$$k(t, s) = \sum_{n, m=-\infty}^{\infty} \underline{\phi}'(t-nT) \underline{A}_{n-m} \underline{\phi}^*(s-mT)$$

(where $\underline{\phi}$ is an M -vector, \underline{A}_{n-m} an $M \times M$ matrix), can be reduced to an operator equation of the form

$$[\underline{A}(f) + \lambda \underline{I}] \cdot \underline{B}(f) = \underline{G}(f) \quad (3-48)$$

where \underline{A} , \underline{B} , \underline{G} are $M \times M$ matrices of periodic functions:

$$\underline{A}(f) = \sum_{r=-\infty}^{\infty} \underline{A}_r e^{-j2\pi r T f}; \quad (3-49)$$

and that the solution to Eq. (3-47) can be obtained from the solution to Eq. (3-48):

$$\underline{B}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1} \underline{G}(f), \quad (3-50)$$

when the indicated matrix-inverse exists.

We now present three types of correlation matrices $\underline{A}(f)$ where this inverse not only exists, but also is relatively easy to compute--even for $M = \infty$. First, however, we mention that since \underline{A} , \underline{G} (and therefore \underline{B}) are $1/T$ -periodic, then the matrix inversion need only be carried out for values of $f \in [-1/2T, 1/2T]$.

(1) Infinite-dimensional diagonal matrices: We begin with a theorem.

THEOREM(3-3): If $\underline{A}(f)$ is the correlation matrix for the kernel k (defined by Eqs. (3-43,44) of Theorem (3-2)), and if k has the wide-sense Markov-like property:

$$k(t+rT, \tau) = \alpha_r k(t, \tau) \quad \forall t, \tau \in [0, T], \quad \forall r$$

for some sequence $\{\alpha_r\}$, then $\underline{A}(f)$ decomposes into the product of a $1/T$ -periodic function and a constant matrix

$$\underline{A}(f) = \underline{A}_0 \sum_r \alpha_r e^{-j2\pi r T f},$$

and if the CON set of basis functions $\{\phi_p\}$ employed in the TSR for k are the eigenfunctions of the operator \underline{A} with kernel k , then \underline{A}_0 is the diagonal matrix \underline{A} whose non-zero elements are the corresponding eigenvalues.

Proof: By definition,

$$\begin{aligned} A_{pq}(f) &\triangleq \sum_r e^{-j2\pi r T f} \int_0^T \int_0^T k(t+rT, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau \\ &= \sum_r \alpha_r e^{-j2\pi r T f} \int_0^T \int_0^T k(t, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau \\ &= A_0^{pq} \sum_r \alpha_r e^{-j2\pi r T f}, \end{aligned}$$

and if $\{\phi_p\}$ are the eigenfunctions corresponding to k , then

$$\begin{aligned} A_0^{pq} &= \int_0^T \int_0^T k(t, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau \\ &= \lambda_q \int_0^T \phi_p^*(t) \phi_q(t) dt \\ &= \lambda_q \delta_{pq}. \end{aligned}$$

QED

There are several CS random processes, of interest to us in this thesis, whose autocorrelation functions k satisfy the hypothesis of this theorem. For the sake of continuity, we give as two examples those processes for which the eigenfunctions were obtained in Subsection 2d:

(1a) Random video signal: From our model developed in Chapter II (Eqs. (2-23;24,25)) we have

$$\begin{aligned} k(t+rT, \tau) &= \alpha_r k(t, \tau) & \forall t, \tau \in [0, T] \\ &= \alpha_r \rho_1^{|t-\tau|} \end{aligned}$$

where

$$\alpha_r = \rho_2^{|(r) \bmod L|} \rho_3^{|r| - |r| \bmod L}$$

and ρ_2, ρ_3 are the line-to-line and frame-to-frame correlation factors, and L is the number of lines per frame. So, from Theorem (3-3) we have the inverse matrix

$$\underline{D}(f) \triangleq [\underline{A}(f) + \lambda \underline{I}]^{-1}$$

with elements

$$d_{pq} = \frac{\delta_{pq}}{\lambda + \lambda \sum_{p_r} \alpha_r e^{-j2\pi r T f}}, \quad (3-51)$$

where the $\{\lambda_p\}$ are given by Eq. (3-27), and if we ignore frame-to-frame correlation ($\rho_3 = 1, L \rightarrow \infty$), then $\alpha_r = \rho_2^{|r|}$ and

$$\sum_{r=-\infty}^{\infty} \alpha_r e^{-j2\pi r T f} = \frac{1 - \rho_2^2}{(1 - \rho_2)^2 + 4\rho_2 \sin^2(\pi T f)}. \quad (3-52)$$

(1b) Time-division-multiplexed signal: As discussed in Subsection 2d it is most convenient to obtain the TSR for the TDM₁ signal in terms of M Karhunen-Loeve expansions--one for each component signal being multiplexed. Thus, the composite correlation matrix is composed of M matrices (the M² - M off-diagonal matrices are zero because the M signals are assumed uncorrelated)

$$\underline{A}(f) = \begin{vmatrix} \underline{A}^1(f) & & 0 \\ & \underline{A}^2(f) & \\ 0 & & \underline{A}^M(f) \end{vmatrix}$$

with elements

$$\underline{A}^l(f) = \sum_r \underline{A}_{r,l} e^{-j2\pi r T f}$$

where

$$A_{r,l}^{pq} = \int_0^{T/M} \int_0^{T/M} k_l(t+rT-\tau) \phi_p^l(t+(l-1)T/M) \phi_q^l(\tau+(l-1)T/M) dt d\tau.$$

But, for the class of correlation functions discussed in Subsection

2d: namely

$$k_l(t-\tau) = \begin{cases} \frac{k_l(0)}{T_l^n} [(t-\tau+T_l)^n u(\tau-t) + (\tau-t+T_l)^n u(t-\tau)], & |t-\tau| \leq T_l, \text{ } n \text{ odd} \\ 0, & |t-\tau| \geq T_l, \end{cases}$$

we have, for $T/2M \leq T_l \leq T(1-1/M)$,

$$k_l(t+rT-\tau) = \alpha_r k_l(t-\tau) \quad \forall t, \tau \in [0, T/M]$$

where

$$\alpha_r = \begin{cases} 1, & r = 0 \\ 0, & r \neq 0 \end{cases}.$$

Hence,

$$A_{r,l}^{pq} = \lambda_p^a \delta_{pq},$$

and

$$\underline{A}^l(f) = \underline{A}^l,$$

and

$$\underline{A}(f) = \begin{vmatrix} \underline{A}^1 & & \\ & \ddots & \\ 0 & & \underline{A}^M \end{vmatrix},$$

so that the inverse matrix given by

$$\underline{D}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1}$$

$$= \begin{vmatrix} \underline{D}^1 & & \\ & \ddots & \\ 0 & & \underline{D}^M \end{vmatrix}$$

has elements

$$d_{pq}^l = \frac{\delta_{pq}}{\lambda + \lambda_p^l}, \quad (3-53)$$

where the $\{\lambda_p^l\}$ are given by Eq. (3-30).

(2) Infinite dimensional diagonal-plus-finite-rank matrices: Again, we begin with a theorem.

THEOREM(3-4): If the inverse matrices $\underline{D} = [\underline{Q} + \underline{RS}']^{-1}$, \underline{Q}^{-1} , and $\underline{F}^{-1} = [\underline{I} + \underline{S}'\underline{Q}^{-1}\underline{R}]^{-1}$ all exist, then

$$\underline{D} = \underline{Q}^{-1} - \underline{Q}^{-1}\underline{R}\underline{F}^{-1}\underline{S}'\underline{Q}^{-1}. \quad (3-54)$$

A proof of this theorem (known as the Matrix Inversion Lemma) can be found in the appendix of reference [49].

Our application of this theorem is to the case where the infinite-dimensional correlation matrix $\underline{A}(f)$ can be decomposed into the sum of an invertible diagonal matrix--call it $\underline{\Lambda}$ --and a finite rank (say N) matrix \underline{RS}' , where $\underline{R}, \underline{S}$ are infinite-by-N dimensional matrices. For then the infinite-dimensional matrix inverse $\underline{D}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1}$ can be computed in terms of the inverse of the N x N matrix $\underline{F}(f) = [\underline{I} + \underline{S}'(f) \cdot [\underline{\Lambda} + \lambda \underline{I}]^{-1} \underline{R}(f)]$ by using Eq. (3-54) of Theorem (3-4) where $\underline{Q} = \underline{\Lambda} + \lambda \underline{I}$.

We present a major class of CS processes whose infinite dimensional correlation matrices can be so decomposed, and then we present a variation of that class:

(2a) CS processes which are the outputs of periodically time-varying finite-dimensional linear dynamical systems driven by white noise:

Since the class of processes which are generated by passing white noise through a finite-dimensional (order) linear dynamical system are of much interest in statistical communication and control theories [49-51]

(because of their analytical and computational tractability), we will present our result on decomposition in the form of a theorem and proof which will be convenient for future reference:

THEOREM(3-5): If x is the T-CS output process of a T-periodic N^{th} order linear dynamical system driven by a vector (of arbitrary dimension) of jointly white WSS processes, then the infinite dimensional correlation matrix (Eqs. (3-43,44)), corresponding to the TSR of Theorem (3-1), for x admits the decomposition:

$$\underline{A}(f) = \underline{A}_0 + \underline{R}(f)\underline{S}' \quad (3-55)$$

where $\underline{R}(f)$, \underline{S} have dimensions $\infty \times 2N$, and if the set of basis functions $\{\phi_p\}$ employed in the TSR for the autocorrelation function k_{xx} are the eigenfunctions of the operator \underline{A} whose kernel is k_{xx} , then \underline{A}_0 is the diagonal matrix \underline{A} whose non-zero elements are the corresponding eigenvalues.

Proof: Consider the state-equation description of an N^{th} order linear dynamical system:

$$\frac{d}{dt}[\underline{y}(t)] = \underline{F}(t)\underline{y}(t) + \underline{b}(t)\underline{n}(t) \quad (3-56)$$

$$x(t) = \underline{c}'(t)\underline{y}(t),$$

where \underline{y} is the $N \times 1$ state vector of random processes, \underline{F} is the $N \times N$ state transition matrix, \underline{b} is the input matrix, \underline{n} is the input vector of jointly white WSS processes with autocorrelation matrix

$$E\{\underline{n}(t)\underline{n}'(\tau)\} = \underline{R}\delta(t-\tau), \quad (3-57)$$

and $\underline{c}'(t)$ is the $1 \times N$ output matrix, and x the output process. Now, the steady-state output x is given explicitly by

$$x(t) = \int_{-\infty}^t \underline{c}'(t) \underline{\phi}(t, \tau) \underline{b}(\tau) \underline{n}(\tau) d\tau \quad (3-58)$$

where $\underline{\phi}$ is the fundamental solution matrix corresponding to the transition matrix \underline{F} [52]. From this expression and Eq. (3-57), it is easily shown that the steady-state autocorrelation function for x is given by

$$k_{xx}(t, s) = \underline{g}'(t) \underline{h}(s) u(t-s) + \underline{g}'(s) \underline{h}(t) u(s-t) \quad (3-59)$$

where the $N \times 1$ matrices g, h are given by

$$\begin{aligned} \underline{g}(t) &= \underline{\phi}(t, t_0) \underline{c}(t) \\ \underline{h}(t) &= \int_{-\infty}^t \underline{\phi}(t_0, \tau) \underline{b}(\tau) \underline{R} \underline{b}'(\tau) \underline{\phi}'(t, \tau) d\tau \underline{c}(t) \end{aligned} \quad (3-60)$$

and t_0 is arbitrary (we've used the property of all fundamental solution matrices: $\underline{\phi}(t, \tau) = \underline{\phi}(t, t_0) \underline{\phi}(t_0, \tau) \quad \forall t_0, t, \tau$).

Thus, using Eq. (3-59) in Eq. (3-44) yields the following expressions for the elements of the matrix of correlation sequences:

$$A_r^{pq} = \begin{cases} \sum_{i=1}^N g_{pi}^*(r) h_{qi}(0), & r > 0 \\ A_0^{pq}, & r = 0 \\ \sum_{i=1}^N h_{pi}^*(r) g_{qi}(0), & r < 0 \end{cases} \quad (3-61)$$

where

$$\begin{aligned} g_{pi}(r) &\triangleq \int_0^T g_i(t+rT) \phi_p(t) dt \\ h_{pi}(r) &\triangleq \int_0^T h_i(t+rT) \phi_p(t) dt, \end{aligned}$$

and g_i, h_i are the elements of the g, h vectors of Eq. (3-60).

Now, substituting Eq. (3-61) into Eq. (3-43), we obtain the following expression for the correlation matrix:

$$\underline{A}(f) = \underline{A}_0 + \underline{G}^*(f)\underline{H}'_0 + \underline{H}^*(f)\underline{G}'_0, \quad (3-62)$$

where the elements of the $M \times N$ dimensional matrices $\underline{G}(f), \underline{H}(f)$ are

$$G_{pi}(f) = \sum_{r=1}^M g_{pi}(r) e^{j2\pi r T f}$$

$$H_{pi}(f) = \sum_{r=1}^M h_{pi}(-r) e^{-j2\pi r T f}$$

and the elements of the $M \times N$ dimensional matrices $\underline{G}_0, \underline{H}_0$ are $\{g_{pi}(0)\}, \{h_{pi}(0)\}$. Hence,

$$\underline{A}(f) = \underline{A}_0 + \underline{R}(f)\underline{S}' \quad (3-63)$$

where

$$\underline{R}(f) \triangleq [\underline{G}^*(f), \underline{H}^*(f)]$$

and

$$\underline{S} \triangleq [\underline{H}_0, \underline{G}_0].$$

Finally, if $\{\phi_p\}$ are the eigenfunctions for the operator with kernel k_{xx} , then as in Theorem (3-3), $\underline{A}_0 = \underline{A}$.

QED

(2b) TDM₁ processes with component signals which have been generated by finite-dimensional time-invariant systems: As a variation on the above result, consider the TDM₁ process discussed throughout this section--but this time consider component processes which can be modeled as the outputs of time-invariant finite-dimensional linear dynamical systems (of order N). As in the previous subsection, the composite correlation matrix takes the form

$$\underline{A}(f) = \begin{bmatrix} \underline{A}^1(f) & & 0 \\ & \underline{A}^2(f) & \\ 0 & & \ddots \\ & & & \underline{A}^M(f) \end{bmatrix}$$

where $\underline{A}^p(f)$ is the correlation matrix for the p^{th} component process, and by Theorem (3-5) can be decomposed as follows

$$\underline{A}^p(f) = \underline{A}_{0,p} + \underline{R}_p(f) \underline{S}_p'$$

where \underline{R}_p , \underline{S}_p are $\infty \times 2N$ dimensional matrices, and if the basis functions used in the TSR are the eigenfunctions, then $\underline{A}_{0,p} = \underline{\Lambda}^p$, the diagonal matrix of eigenvalues. Now, since

$$[\underline{A}(f) + \lambda \underline{I}]^{-1} = \begin{bmatrix} [\underline{A}^1(f) + \lambda \underline{I}]^{-1} & & 0 \\ & [\underline{A}^2(f) + \lambda \underline{I}]^{-1} & \\ 0 & & \ddots \\ & & & [\underline{A}^M(f) + \lambda \underline{I}]^{-1} \end{bmatrix} \quad (3-64)$$

then this composite inverse can be obtained by employing Theorem (3-4) to obtain the inverse of each of the M component matrices--the total computation requiring the inversion of M matrices of dimension $2N \times 2N$.

(3) Finite-dimensional diagonal and diagonal-plus-rank-1 matrices:

Our third and last class of correlation matrices to be considered is, in fact, a special case of the previous two classes. We consider here correlation matrices which are not only finite-dimensional, but also either diagonal or diagonal-plus-rank-1.

If $\underline{A}(f) = \underline{\Lambda}(f)$, a diagonal matrix with non-zero elements $\{\lambda_p(f)\}$, then the inverse matrix

$$\underline{D}(f) = [\underline{A}(f) + \lambda \underline{I}]^{-1}$$

has elements

$$d_{pq}(f) = \frac{\delta_{pq}}{\lambda_p(f) + \lambda}, \quad (3-65)$$

or if $\underline{A}(f) = \underline{\Lambda}(f) + \underline{R}(f)\underline{S}'(f)$, where $\underline{R}(f)$, $\underline{S}(f)$ are $M \times 1$ vectors with elements $\{r_p(f)\}, \{s_p(f)\}$, then, from Theorem (3-4), the inverse matrix $\underline{D}(f)$ has elements

$$d_{pq}(f) = \frac{\delta_{pq}}{\lambda_p(f) + \lambda} - \frac{\left[\frac{r_p(f)}{\lambda_p(f) + \lambda} \right] \left[\frac{s_q(f)}{\lambda_q(f) + \lambda} \right]}{\left[1 + \sum_{i=1}^M \frac{s_i(f)r_i(f)}{\lambda_i(f) + \lambda} \right]}, \quad (3-66)$$

so that, in both cases, the inverse matrix $\underline{D}(f)$ is given explicitly with no need for further computation.

Now, the purpose here is to present two specific examples of CS processes--one of which has a finite-dimensional diagonal correlation matrix, and the other of which has a finite-dimensional diagonal-plus-rank-1 matrix. Both of the examples given here are models of random signals which are of much interest in communications:

(3a) Time-division-multiplexed PAM signals: As an alternative to the TDM_1 and TDM_2 schemes introduced in Chapter II for time-multiplexing a multiplicity of signals, we consider here the TDM-PAM scheme. If our M component processes $\{y_p\}$ are bandlimited to $[-1/2T, 1/2T]$, then they can be periodically sampled every T seconds without loss of information. If these M sample sequences are now interleaved in time to provide a single composite sequence of samples--one every T/M seconds--then this composite sequence can be employed to amplitude-modulate a stream of pulses to obtain a time-division-multiplexed PAM signal x :

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M y_p(nT) q(t - nT - pT/M). \quad (3-67)$$

If the M component processes are jointly WSS, then x is a T -CS process with an M^{th} order TSR with a matrix of correlation sequences given by

$$A_{n-m}^{pq} = E\{y_p(nT) y_q(mT)\}$$

so that, if the component processes are also mutually uncorrelated, then

$$A_{n-m}^{pq} = A_{n-m}^{pp} \delta_{pq},$$

and the correlation matrix $\underline{A}(f)$ is diagonal:

$$A_{pq}(f) = \delta_{pq} \sum_r A_r^{pp} e^{-j2\pi r T f}. \quad (3-68)$$

(3b) Synchronous M -ary signals: In Subsection 2e, we pointed out that synchronous M -ary signals such as FSK admit M^{th} order TSR's where the matrix of correlation sequences is a matrix of joint probabilities:

$$A_{n-m}^{pq} = \text{Prob}[b_n = \alpha_p \text{ and } b_m = \alpha_q].$$

Now, if we assume that the random variables in the modulating sequence $\{b_n\}$ are mutually statistically independent, then

$$A_{n-m}^{pq} = \begin{cases} \text{Prob}[b_n = \alpha_p] \text{Prob}[b_m = \alpha_q], & n \neq m \\ \delta_{pq} \text{Prob}[b_n = \alpha_p], & n = m \end{cases}$$

$$= P(p)P(q)(1 - \delta_{nm}) + \delta_{pq}P(p)\delta_{nm}$$

where $P(p) \triangleq \text{Prob}[b_n = \alpha_p]$ for all n . Hence, the $M \times M$ correlation matrix is given by

$$A_{pq}(f) = \delta_{pq}P(p) + P(p)P(q) \sum_{r \neq 0} e^{-j2\pi r T f}, \quad (3-69)$$

and is composed of the sum of a diagonal matrix and a rank-1 matrix.

g) Implementation of periodically time-varying linear systems.

As shown in the previous subsection, the TSR of an autocorrelation function which is the kernel of a linear integral transformation (operator) is reflected in the kernel of another transformation which is the solution of an operator equation (Eqs. (3-38,47)). For some applications (especially estimation and detection as discussed in the following Chapter) it is desired to physically implement the solution operator as a linear system whose impulse response function is the kernel.

An impulse response function h which has an M^{th} order TSR

$$h(t, \tau) = \sum_{n, m=-\infty}^{\infty} \sum_{p, q=1}^M B_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(\tau-mT)$$

has the straightforward implementation shown in Fig. (3-9), where the input is applied to a bank of M time-invariant filters (with transfer functions $\{\phi_p^*\}$) whose outputs are periodically sampled every T seconds and applied to an $M \times M$ matrix of time-invariant sampled-data filters [53,54] (with transfer functions $\{B_{pq}(f)\}$ which are z transforms of the impulse response sequences $\{B_n^{pq}\}$). The sequences of output samples from the sampled-data filters (assumed weighted sequences of impulses) are then applied to a bank of M pulse-generating, time-invariant filters (with transfer functions $\{\phi_p\}$) whose outputs are summed to form the final output of the system.

Note that this system can be obtained from the resolution-reconstruction system of Fig. (3-7) simply by inserting the matrix of sampled-data filters between the resolution portion and the reconstruction portion. Similarly, a multiplier-integrator type of implementation for the system with impulse response function h can be obtained from the resolution-reconstruction system of Fig. (3-5) by, again, inserting the matrix of sampled-data filters between the resolution and reconstruction portions.

Finally, it should be mentioned that the impulse response function h will not, in general, represent a causal system, so that a physical implementation must incorporate a delay, say t_0 , so that $\tilde{h}(t, \tau) = h(t - t_0, \tau)$ is the impulse response which is implemented. Notice, however, that if one started with a causal h :

$$h(t, \tau) = h(t, \tau)u(t - \tau)$$

where u is the unit step function, and one expanded it into a TSR, then

the sampled-data-filter impulse response sequences would satisfy

$$B_{n-m}^{pq} = \int_0^T \int_0^T h(t+(n-m)T, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau$$

$$= 0 \quad \text{for } n-m < 0.$$

Hence, the sampled-data filters would be causal. Furthermore since the output filters $\{\phi_p\}$ are causal, and the input filters $\{\phi_p^*\}$ can be made causal with a delay of length T , then the overall (approximate) realization (which uses only a finite number of terms) shown in Fig. (3-9) would be causal with a delay of $t_0 = T$. Note that a finite order TSR cannot be causal without the incorporation of at least a T -second delay.

h) Generalized Fourier transforms for cyclostationary processes.

A rather disappointing quirk of the transfer function theory for linear time-invariant systems (as it stands) is that it applies only to deterministic signals, so that random input and output signals cannot, in general, be related by the simple formula

$$X(f) = H(f)Y(f)$$

where X , Y , H are, respectively, the Fourier Transforms of the output x , the input y , and the impulse response function h . The reason being that Fourier transforms of random signals are not, in general, defined. The standard argument against the possible existence of a Fourier transform of a stationary random process is: the Fourier transform--like any other linear transformation of a random process--should be interpreted as the ensemble of Fourier transforms of the deterministic functions which comprise the ensemble of realizations of the process; but due to

the "continuing" nature of a stationary process, its realizations will not be in $L^2(-\infty, \infty)$, and therefore will not be Fourier transformable. The prevalence of this attitude is evidenced by this statement from a popular modern engineering textbook on random processes [Ref. 3, pp. 336,465]: "We assume the reader is familiar with linear systems and transform techniques. We should emphasize, however, that we transform only deterministic signals and not stochastic processes.... Stationary processes have, in general, no Fourier transforms."

The purpose of this subsection is to employ the TSR of Theorem (3-1) to define a simple, yet adequate, generalized Fourier transform for the class of all finite mean-square CS (and WSS) processes. With this generalized Fourier transform, the transfer function methods, which are so useful in the analysis of linear time-invariant systems and deterministic signals, can be directly extended to the analysis of these systems and random signals. For example, the frequency-domain input-output relations of Chapter II (Eqs. (2-7,8,9)), which were derived in a round-about way in the time domain, can be re-derived in the frequency domain in one or two simple steps. This is, in fact, done in [Ref. 3, pp. 465-467] by "...ignoring the mathematical difficulties."

We begin by presenting a generalized version of the well known formal derivation of the generalized Fourier transform of a periodic function. Consider any periodic function x in the Hilbert space $L^2_{P(T)}(-\infty, \infty)$, defined in Sec. (3-1). If $\{\phi_p\}$ is any CON set in $L^2[0,T]$, then x admits the deterministic translation series representation

$$\begin{aligned}
 x(t) &= \sum_{p=1}^{\infty} \sum_{n=-\infty}^{\infty} a_{np} \phi_p(t-nT) \quad \text{a.e. } t \in (-\infty, \infty) \\
 &= \sum_p a_{0p} \sum_n \phi_p(t-nT)
 \end{aligned}$$

where

$$a_{np} = \int_0^T x(t+nT) \phi_p^*(t) dt = a_{0p} \quad \forall n.$$

Now, we define the generalized Fourier transform for any

$x \in L^2_{p(T)}(-\infty, \infty)$ to be the equivalence class of transformations $F_{L^2_p}\{\cdot\}$

with representatives

$$\begin{aligned}
 F_{L^2_p}\{x(t)\} &= X(f) \triangleq \sum_p a_{0p} \sum_n F_{L^2}\{\phi_p(t-nT)\} \\
 &= \sum_p a_{0p} \sum_n \phi_p(f) e^{-j2\pi nTf} \\
 &= \frac{1}{T} \sum_p a_{0p} \sum_m \phi_p(m/T) \delta(f-m/T)
 \end{aligned}$$

where $F_{L^2}\{\cdot\}$ is the usual Fourier transform defined on $L^2(-\infty, \infty)$, and where the last equation is the Poisson sum formula [2,55]. The equivalence relation which generates this equivalence class is

$$x_1 \sim x_2 \Leftrightarrow \int_{-\infty}^{\infty} |x_1(f) - x_2(f)| df = 0,$$

where x_1 and x_2 are the two transforms of x corresponding to the two representatives of $F_{L^2_p}$ obtained from any two CON sets in $L^2[0, T]$, and x is any function in $L^2_{p(T)}$. It is not difficult to show that all representatives of $F_{L^2_p}\{\cdot\}$ corresponding to any specific signal x , do indeed satisfy this equivalence condition.

Now consider the specific CON set of complex exponentials

$$\phi_p(t) = \frac{1}{\sqrt{T}} e^{j2\pi pt/T} w(t)$$

where w is the unit gate function on $[0, T]$. The Fourier transforms of the $\{\phi_p\}$ are

$$\phi_p(f) = \frac{1}{\sqrt{T}} W(f - p/T)$$

where

$$W(f) = \frac{\sin(\pi f T)}{\pi f}$$

so that

$$\phi_p(nT) = \sqrt{T} \delta_{pn}.$$

Also, the TSR coefficients are

$$\sqrt{T} a_{0p} = \int_0^T x(t) e^{-j2\pi pt/T} dt \triangleq a_p$$

and are the familiar Fourier coefficients for the harmonic Fourier series representation for x . Thus, our generalized Fourier transform now takes the well known form [55]

$$X(f) = \sum_{p=-\infty}^{\infty} a_p \delta(f - p/T).$$

This generalized Fourier transform can easily be extended from deterministic to random periodic functions, as discussed in [Ref. 3, p. 467], simply by allowing the Fourier coefficients to be random variables. In fact, our more general definition of the generalized Fourier transform $F_{L^2_p}$ can just as easily be extended from deterministic

periodic functions to cyclostationary random processes:

DEFINITION: The generalized Fourier transform for any random process, say x , which is contained in the Hilbert space $H_{CS(T)}^{(-\infty, \infty)}$ is defined to be the equivalence class of transformations $F_{H_{CS}}\{\cdot\}$ with representatives

$$\begin{aligned} F_{H_{CS}}\{x(t)\} &= X(f) \triangleq \sum_{p=1}^{\infty} \sum_{n=-\infty}^{\infty} a_{np} F_{L^2}(\phi_p(t-nT)) \\ &= \sum_p \sum_n a_{np} \phi_p(f) e^{-j2\pi n T f}, \end{aligned}$$

where $\{\phi_p\}$ is any CON set in $L^2[0, T]$ (extended to the zero-function outside $[0, T]$) and $\{a_{np}\}$ are the random TSR representors

$$a_{np} = \int_0^T x(t+nT) \phi_p^*(t) dt$$

and $F_{L^2}\{\cdot\}$ is the usual Fourier transform defined on $L^2(-\infty, \infty)$.

The equivalence relation which generates this equivalence class is

$$X_1 \sim X_2 \Leftrightarrow \iint_{-\infty}^{\infty} |E\{(X_1(f) - X_2(f))(X_1(v) - X_2(v))^*\}| df dv = 0$$

where X_1 and X_2 are the two transforms of x corresponding to the two representatives of $F_{H_{CS}}$ obtained from any two CON sets in $L^2[0, T]$, and x is any random process in $H_{CS(T)}^{(-\infty, \infty)}$.

That this relation is indeed an equivalence relation [2] can be verified by showing

$$\iint_{-\infty}^{\infty} |E\{(X_1(f) - X_2(f))(X_1(v) - X_2(v))^*\}| df dv = 0$$

$$\Leftrightarrow E\{|x_1(t) - x_2(t)|^2\} = 0 \quad \forall t \in (-\infty, \infty),$$

where x_1 and x_2 are the random processes with autocorrelation functions

$$k_{x_1 x_1}(t, s) = F \cdot F_{L_p^2}^{-1} \{E\{X_1(f) X_1^*(v)\}\} \quad i = 1, 2$$

where $F \cdot F_{L_p^2}^{-1}$ is the generalized double Fourier transform inverse for periodic functions of two variables; and using the fact that

$$x_1 \sim x_2 \Leftrightarrow E\{|x_1(t) - x_2(t)|^2\} = 0 \quad \forall t \in (-\infty, \infty)$$

is a valid equivalence relation.

The following theorem establishes the equivalence under expectation of our generalized Fourier transform for CS processes and our generalized Fourier transform for deterministic periodic functions:

THEOREM(3-6): The generalized Fourier transforms $F_{H_{CS}}$, $F_{L_p^2}$ are equivalent under expectation in the following sense:

$$E\{F_{H_{CS}}\{x(t)\}\} = F_{L_p^2}\{E\{x(t)\}\} = F_{L_p^2}\{m_x(t)\}$$

$$\begin{aligned} E\{F_{H_{CS}}\{x(t)\} F_{H_{CS}}^*\{x(s)\}\} &= F \cdot F_{L_p^2}\{E\{x(t) x^*(s)\}\} \\ &= F \cdot F_{L_p^2}\{k_{xx}(t, s)\} = K_{xx}(\cdot, \cdot) \end{aligned}$$

for every $x \in H_{CS(T)}(-\infty, \infty)$.

Proof: Let $X(f)$ be the representative of the equivalence class

$F_{H_{CS}}\{\cdot\}$ corresponding to any TSR, say $\{\phi_p, \{a_{np}\}\}$.

Then:

$$\begin{aligned}
(i) \quad E\{X(f)\} &= E\left\{\sum_p \sum_n a_{np} F_{L^2}\{\phi_p(t-nT)\}\right\} \quad (\text{by definition}) \\
&= E\left\{\sum_p \sum_n \int_0^T x(\tau+nT) \phi_p^*(\tau) d\tau F_{L^2}\{\phi_p(t-nT)\}\right\} \quad (\text{Eq. (3-12)}) \\
&= \sum_p \sum_n \int_0^T E\{x(\tau+nT)\} \phi_p^*(\tau) d\tau F_{L^2}\{\phi_p(t-nT)\} \quad (\text{assumed valid}) \\
&= \sum_p \sum_n \int_0^T m_x(\tau+nT) \phi_p^*(\tau) d\tau F_{L^2}\{\phi_p(t-nT)\} \quad (\text{by def.}) \\
&= F_{L^2}\{m_x(t)\} \quad (\text{by def.}) \\
&= \frac{1}{T} \sum_p \int_0^T m_x(\tau) \phi_p^*(\tau) d\tau \sum_m \phi_p(m/T) \delta(f-m/T);
\end{aligned}$$

$$\begin{aligned}
(ii) \quad E\{X(f)X^*(v)\} &= E\left\{\sum_{pq} \sum_{nm} a_{np} a_{mq}^* F_{L^2}\{\phi_p(t-nT)\} F_{L^2}\{\phi_q^*(s-mT)\}\right\} \quad (\text{by Def.}) \\
&= \sum_{pq} \sum_{nm} E\{a_{np} a_{mq}^*\} F_{L^2}\{\phi_p(t-nT)\} F_{L^2}\{\phi_q^*(s-mT)\} \\
&\quad (\text{assumed valid}) \\
&= \sum_{pq} \sum_{nm} \Lambda_{n-m}^{pq} F_{L^2}\{\phi_p(t-nT)\} \phi_q^*(s-mT) \quad (\text{by def. and Eq. (3-3)}) \\
&= F_{L^2}\{k_{xx}(t,s)\} \quad (\text{by def. and Eq. (3-2)}) \\
&= k_{xx}(f,v) \quad (\text{by def.}) \\
&= \frac{1}{T} \sum_{pq} \Lambda_{pq}(v) \phi_p(f) \phi_q^*(v) \sum_m \delta(f-v+m/T) \quad (\text{Eq. (3-4)}).
\end{aligned}$$

QED

It should be mentioned that this theorem is, in a sense, the extension of Helly's theorem (Ref. [35], p.247)--as applied to time averages of ergodic stationary processes--to ensemble averages of non-ergodic cyclostationary (and stationary) processes.

Note that the expected singular behavior of $F_{H_{CS}}\{x\}$ is, in more than one way, "expected" since it shows up in the mean function $E\{F_{H_{CS}}\{x(t)\}\}$, as a string of impulses, and in the autocorrelation function $E\{F_{H_{CS}}\{x(t)\}F_{H_{CS}}^*\{x(s)\}\}$, as a string of impulse fences.

Note also that since, for WSS processes, we have

$$K_{xx}(f, \nu) = K_{xx}(f)\delta(f-\nu)$$

where $K_{xx}(f)$ is the power spectral density for the process x , then we now have two new formulas for the power spectral density:

$$K_{xx}(f) = \int_{-\infty}^{\infty} E\{X(f)X^*(\nu)\}d\nu = E\{X(f)x(0)\}$$

where

$$X(f) = F_{H_{CS}}\{x(t)\}.$$

The first topic considered in the next section provides a concrete example of the utility of a generalized Fourier transform for CS random processes.

3. Harmonic Series Representation

a) Definition. In this section we develop a particularly interesting and unique translation series representation which is intimately related to the harmonic TSR discussed in Sec. 2d, and is, in fact, the frequency-time dual of the harmonic TSR. This duality is immediately brought to evidence if we introduce the representation in the following manner: We begin by representing the Fourier transform $X(f)$ of the T-CS process $x(t)$ with the frequency-domain harmonic translation series

$$X(f) = \sum_n \sum_p a_{np} \theta_n(f-p/T), \quad (3-70)$$

where $\{\theta_n\}$ is the CON (in $L^2[-1/2T, 1/2T]$) set of complex exponentials

$$\theta_n(f) = \sqrt{T} e^{-j2\pi nTf} W(f) \quad (3-71)$$

where W is the gate function defined by

$$W(f) = \begin{cases} 1, & f \in [-1/2T, 1/2T] \\ 0, & f \notin [-1/2T, 1/2T] \end{cases} \quad (3-72)$$

and where the representors $\{a_{np}\}$ are given by

$$a_{np} = \int_{-1/2T}^{1/2T} X(f+p/T) \theta_n^*(f) df. \quad (3-73)$$

Now, if we take the inverse Fourier transform of this frequency-domain TSR, then we obtain the following TSR in the time-domain:

$$x(t) = \sum_n \sum_p a_{np} \phi_p(t-nT) \quad (3-74)$$

where

$$\phi_p(t) = \sqrt{T} e^{j2\pi pt/T} w(t) \quad (3-75)$$

and w is the inverse Fourier transform of the gate function W

$$w(t) = \frac{\sin(\pi t/T)}{\pi t} \quad (3-76)$$

Using Parseval's relation [2] in Eq. (3-73) yields the time-domain formulas for the representors:

$$a_{rp} = \int_{-\infty}^{\infty} x(t+nT) \phi_p^*(t) dt. \quad (3-77)$$

Thus, we see that our frequency-time dual of the harmonic TSR of Section 2d is a TSR in both the time- and frequency domains.

Now, analogous to our development of representations in terms of PAM processes in Section 2b, we collect terms in our time-domain TSR of Eq. (3-74) to obtain

$$x(t) = \sum_p a_p(t) e^{j2\pi pt/T} \quad (3-78)$$

where

$$a_p(t) = \sqrt{T} \sum_n a_{np} w(t-nT). \quad (3-79)$$

Employing the formulas for $\{a_{np}\}$ in the above equation yields the more direct formulas for the continuous-time representors:

$$a_p(t) = \int_{-\infty}^{\infty} w(t-\tau) x(\tau) e^{-j2\pi p\tau/T} d\tau. \quad (3-80)$$

We take Eqs. (3-78,80) as our formal definition of the harmonic series representation.²¹ Note that, in terms of Eq. (3-79), we can interpret $\{a_p(t)\}$ as PAM processes, so that we again see that this harmonic series representation (HSR) is the frequency-time dual of the harmonic TSR of Section 2d--this time in the sense that the continuous-time representors in the harmonic TSR are PAM processes with pulses which are rectangular (width T) in the time domain, and the representors in the HSR are PAM processes with pulses which are rectangular (width $1/T$) in the frequency domain.

For this harmonic series representation, we have the following fundamental theorem:

THEOREM(3-7): Every T-CS process x admits the mean-square equivalent harmonic series representation

$$\begin{aligned}
 x(t) &= \sum_{p=-\infty}^{\infty} a_p(t) e^{j2\pi p t/T} & \text{a.e. } t \in (-\infty, \infty) \\
 a_p(t) &= \int_{-\infty}^{\infty} w(t-\tau) x(\tau) e^{-j2\pi p \tau/T} d\tau \\
 w(t) &= \frac{\sin(\pi t/T)}{\pi t} & (3-81)
 \end{aligned}$$

where the continuous-time representors $\{a_p\}$ are jointly WSS and bandlimited to the interval $[-1/2T, 1/2T]$.

²¹ This representation has previously been presented, but not developed, by Ogura [6]; e.g., the joint wide-sense-stationarity of the representors has not (to my knowledge) been recognized except in our preliminary paper [43], nor has the fundamental relationship of this representation to the class of TSR's been previously discussed.

Proof: Define

$$e(t) \triangleq E\{|x(t) - \sum_p a_p(t)e^{j2\pi pt/T}|^2\}.$$

Expanding the square and interchanging summation and expectation yields

$$\begin{aligned} e(t) &= k_{xx}(t, t) - 2\operatorname{Re}\left[\sum_p E\{a_p^*(t)x(t)\}e^{-j2\pi pt/T}\right] \\ &\quad + \sum_{p,q} E\{a_p^*(t)a_q(t)\}e^{-j2\pi(p-q)t/T}. \end{aligned}$$

Substituting the given expression for a_p and interchanging summation and integration results in

$$\begin{aligned} e(t) &= k_{xx}(t, t) - 2\operatorname{Re}\left[\int_{-\infty}^{\infty} \sum_p w(t-\tau)k_{xx}(t, \tau)e^{-j2\pi p(t-\tau)/T}d\tau\right] \\ &\quad + \iint_{-\infty}^{\infty} \sum_{p,q} w(t-\tau)w(t-\gamma)k_{xx}(\tau, \gamma)e^{-j2\pi p(t-\tau)/T}e^{j2\pi q(t-\gamma)/T}d\tau d\gamma. \end{aligned}$$

But, from the Poisson sum formula [2], we have

$$\begin{aligned} \sum_p w(t-\tau)e^{-j2\pi p(t-\tau)/T} &= T \sum_q w(t-\tau)\delta(t-\tau-qT) \\ &= \delta(t-\tau). \end{aligned}$$

Hence,

$$\begin{aligned} e(t) &= k_{xx}(t, t) - 2\operatorname{Re}[k_{xx}(t, t)] + k_{xx}(t, t) \\ &= 0 \quad \text{e. t. } t \in (-\infty, \infty). \end{aligned}$$

Furthermore, since each process $a_p(t)$ is obtained from the T-CS process $x(t)e^{-j2\pi pt/T}$ via a bandlimiting filter (impulse response function

$\sin(\pi t/T)/\pi t$ with passband $[-1/2T, 1/2T]$, then Theorem (2-3) of Chapter II insures that $a_p(t)$ is a WSS process which is bandlimited to $[-1/2T, 1/2T]$. In fact, since the processes $\{x(t)e^{-j2\pi pt/T}; p = 0, \pm 1, \pm 2, \dots\}$ are jointly T-CS, then the representors $\{a_p(t); p = 0, \pm 1, \pm 2, \dots\}$ are jointly WSS.

QED

The autocorrelation function for every T-CS process has a harmonic series representation which follows directly from the HSR for the process itself:

$$k_{xx}(t, s) = \sum_{p, q} k_{pq}(t-s) e^{j2\pi(pt-qs)/T}, \quad (3-82)$$

where $k(t-s)$ is the correlation matrix for the jointly WSS representors:

$$\begin{aligned} k_{pq}(t-s) &\triangleq E\{a_p(t)a_q^*(s)\} \\ &= \iint_{-\infty}^{\infty} w(t-\tau)w(s-\gamma)k_{xx}(\tau, \gamma)e^{-j2\pi(p\tau-q\gamma)/T} d\tau d\gamma. \end{aligned} \quad (3-83)$$

Similarly, the double Fourier transform of the autocorrelation function has the representation

$$K_{xx}(f, \nu) = \sum_{p, q} K_{pq}(f-p/T)\delta(f-\nu+(q-p)/T), \quad (3-84)$$

where $K(f)$ is the single Fourier transform of the correlation matrix $k(t)$:

$$\begin{aligned} K_{pq}(f) &= F[k_{pq}(t)] \\ &= \int_{-1/2T}^{1/2T} K_{xx}(f+p/T, \nu+q/T) d\nu W(f), \end{aligned} \quad (3-85)$$

and is bandlimited to the interval $[-1/2T, 1/2T]$.

It is interesting to note that the correlation matrix $\underline{K}(f)$ for the HSR of any T-CS process is simply related to the correlation matrix $\underline{A}(f)$ corresponding to any TSR for that process as follows:

$$K_{pq}(f) = \frac{W(f)}{T} \sum_{p', q'} \phi_{p'}(f+p/T) A_{p' q'}(f) \phi_{q'}^*(f+q/T) \quad (3-86)$$

where $\{\phi_p\}$ are the Fourier transforms of the basis functions employed in the TSR. Furthermore, if the TSR is the specific one from which the HSR is derived (Eqs. (3-74)-(3-77)), then

$$\phi_p(f) = \sqrt{T} W(f-p/T) \quad (3-87)$$

and, we have

$$\begin{aligned} K_{pq}(f) &= W(f) A_{pq}(f) \\ A_{pq}(f) &= \sum_{n=-\infty}^{\infty} K_{pq}(f-n/T). \end{aligned} \quad (3-88)$$

Thus, the HSR correlation matrix $\underline{K}(f)$ is just a bandlimited (to $[-1/2T, 1/2T]$) version of the $1/T$ periodic TSR correlation matrix $\underline{A}(f)$; i.e., $\underline{K}(f)$ is one (centered) period of $\underline{A}(f)$.

Now, substituting Eqs. (3-87,88) into Eq. (3-7) for the power spectral density of the stationarized version of a T-CS process, we obtain the HSR formulas:

$$\begin{aligned} K_{xx}(f) &= \sum_{p=-\infty}^{\infty} A_{pp}(f) W(f-p/T) \\ &= \sum_{p=-\infty}^{\infty} K_{pp}(f-p/T) \end{aligned} \quad (3-89)$$

and we see that the stationary character of a CS process is determined by the diagonal elements in the HSR correlation matrix.

b) Resolution and reconstruction. The harmonic series representation is perhaps most transparent when viewed in terms of the process-resolution operation which is defined by the integral equation Eq. (3-81) of Theorem (3-7). Applying the convolution theorem to this equation yields the following simple formula for the Fourier transform of the representor a_p :

$$A_p(f) = X(f + p/T)W(f) \quad (3-90)$$

where W is the gate function defined in Eq. (3-72). Thus, the resolution operation corresponds to equipartitioning the "frequency line" into support intervals of length $1/T$ upon which the process is then decomposed as shown in Fig. (3-10) for a hypothetical realization of the Fourier transform of the process.

This resolution operation has the straightforward implementation shown in Fig. (3-11) where W is the ideal low-pass filter with transfer function $W(f)$. Furthermore, the original process can be reconstructed from the representors simply by shifting them (in the frequency domain) and adding them. The corresponding implementation is also shown in Fig. (3-10).

Note that, paralleling the results in Section 2c, a resolution-reconstruction system which decomposes CS processes into the real and imaginary parts of the representors $\{a_p(t)\}$ can be implemented using only real sinusoidal multiplier functions in the structure shown in Fig. (3-11).

c) Properties of the harmonic series representation. In this subsection, we present several properties of the HSR which illustrate

its convenience for characterizing CS processes.

(1) Stationarity: Since the components $\{a_p(t)e^{j2\pi pt/T}\}$ of the HSR are individually WSS, but not jointly WSS (because of the exponential factor), then a CS process will, in general, be WSS if and only if the various representors $\{a_p\}$ are uncorrelated; i.e., if and only if the correlation matrix is diagonal.

(2) Phase randomization: If \tilde{x} is the process which is derived from a T-CS process x by the introduction of a random phase variable θ :

$\tilde{x}(t) = x(t+\theta)$, then as discussed in Section 5c of Chapter II \tilde{x} is also T-CS. Furthermore, if $p_\theta(\cdot)$ is the PDF for θ and $P_\theta(\cdot)$ is its Fourier transform (conjugate characteristic function), then the elements of the correlation matrix $\tilde{K}(f)$ for \tilde{x} can easily be shown to be related to the elements of the correlation matrix $K(f)$ for x by the formula:

$$\tilde{K}_{pq}(f) = P_\theta((p-q)/T) K_{pq}(f). \quad (3-91)$$

Now, since $p_\theta(\cdot)$ is a PDF, then $P_\theta(r/T) \leq P_\theta(0) = 1$, and we see that the off-diagonal elements are attenuated by the phase-randomization.

Furthermore, if $P_\theta((p-q)/T) = \delta_{pq}$, then $\tilde{K}(f)$ will be diagonal and \tilde{x} will be WSS (a result previously obtained in Theorem (2-15) of Chapter II).

(3) Bandlimitedness: As defined in Section 2a of Chapter II, a random process x is bandlimited to the interval $[-B, B]$ if and only if the double Fourier transform of its autocorrelation function satisfies the bandlimiting constraint:

$$K_{xx}(f, \nu) = 0$$

for $|f| \geq B$ and for $|\nu| \geq B$. Now, using Eq. (3-85) and this definition,

it is easily shown that if x is a T-CS process which is bandlimited to $[-(N+1/2)/T, (N+1/2)/T]$ then x admits an M^{th} order HSR; i.e., all representors with indices greater (in magnitude) than M are zero, and the correlation matrix is $(2M+1)$ dimensional. Similarly, if x is a "band-pass" process:

$$K_{xx}(f, \nu) = 0$$

for $||f| - f_0| \geq B$, and for $||\nu| - f_0| \geq B$, then x admits an HSR with only $2M$ terms (where $B = (N+1/2)/T$).

(4) Time-invariant filtering: If \tilde{x} is the output of a time-invariant filter with T-CS input x , then as discussed in Sec. 2a of Chapter II, \tilde{x} is also T-CS. Furthermore, if G is the transfer function for the filter, then the elements of the correlation matrix $\tilde{K}(f)$ for \tilde{x} can be related to the elements of the correlation matrix $K(f)$ for x by the formula:

$$\tilde{K}_{pq}(f) = W(f)G(f+p/T)G^*(f+q/T)K_{pq}(f). \quad (3-92)$$

Note that if G is an ideal low-pass filter with cutoff frequency f_0 , then all elements in the matrix $\tilde{K}(f)$ with either index greater (in magnitude) than $N \triangleq \min\{n; n > Tf_0 + 1/2\}$ will be zero, so that x can be represented by an M^{th} order HSR--a result which is obvious from the above discussion on bandlimiting.

d) Examples. In this subsection, we present several specific types of CS processes which admit particularly simple finite order HSR's.

(1) Amplitude-modulated signals: Our first example of a CS process introduced in Chapter II (Sec. 2b) was the amplitude-modulated (AM) signal with realizations of the form:

$$x(t) = p(t)y(t)$$

where p is a T -periodic deterministic function and y is a WSS process.

Now if p has the finite harmonic Fourier series representation

$$p(t) = \sum_{q=-M}^M a_q e^{j2\pi q t/T}$$

and y is bandlimited to $[-1/2T, 1/2T]$ then x admits an M^{th} order HSR with representors

$$a_p(t) = \begin{cases} a_p y(t), & p = 0, \pm 1, \pm 2, \dots, \pm M \\ 0, & |p| > M. \end{cases}$$

and if p is a sinusoid (as it most often is in communications applications), then $M = 1$, and the correlation matrix has the form:

$$\underline{K}(f) = K_{yy}(f) \begin{vmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{vmatrix}$$

(2) Frequency-division-multiplexed signals: It was shown in Section 2b of Chapter II that a multiplicity of T -CS AM signals can be added to form a composite T -CS frequency-division-multiplexed (FDM) signal with realizations of the form

$$x(t) = \sum_{p=1}^M y_p(t) \cos(\omega_0 t + 2\pi p t/T)$$

where ω_0 is an integer multiple of $2\pi/T$, and $\{y_p\}$ are each bandlimited to $[-1/2T, 1/2T]$. Now, by expanding the sinusoid into the sum of exponentials, we obtain the HSR

$$x(t) = \sum_{p=-(M+N)}^{(M+N)} a_p(t) e^{j2\pi p t/T},$$

where $N = \omega_0 T/2\pi$, and

$$a_p(t) = \begin{cases} \frac{1}{2} y_p(t), & p = \pm(N+1), \pm(N+2), \dots, \pm(N+M) \\ 0, & N+M < |p| < N. \end{cases}$$

If the components $\{y_p\}$ are uncorrelated, then the composite FDM signal has a sparse correlation matrix of the form:

$$\underline{K} = \begin{vmatrix} K_{(-N-M)(-N-M)} & & & & & K_{(-N-M)(N+M)} \\ & \ddots & & & & \\ & & K_{(-N)(-N)} & & & \\ & & & 0 & & \\ & 0 & & 0 & & 0 \\ & & K_{(N)(-N)} & & & \\ & & & 0 & & \\ & & & & K_{(N)(N)} & \\ & & & & & \ddots \\ K_{(N+M)(-N-M)} & & & & & K_{(N+M)(N+M)} \end{vmatrix}$$

with elements given by

$$K_{pq}(f) = \begin{cases} \frac{1}{4} K_r(f), & |p| = |q| = N + r \\ 0, & \text{otherwise} \end{cases}$$

where $K_r(f)$ is the power spectral density for y_r .

(3) Random periodic signals: In Section 5a of Chapter II, we discussed the random periodic signal as a reasonable model for many biomedical and communications test signals (modeled at the receiver). These signals can be expressed in terms of a Fourier series

$$x(t) = \sum_{p=-M}^M a_p e^{j2\pi pt/T}$$

where $\{a_p\}$ are random Fourier coefficients. Clearly then, these signals admit M^{th} order HSR's with representors $a_p(t) = a_p$, so that the correlation matrix $\underline{K}(f)$ factors into the product of a constant matrix and an impulse function.

e) Solution of integral equations and realization of periodically time-varying systems. As discussed in Section 2f, the TSR for an autocorrelation function can be employed to convert an integral equation to a matrix equation which can be solved by matrix inversion. Now, since the HSR is derived from a TSR, then it too can be similarly employed. In fact, using Eqs. (3-75), (3-88) in Theorem (3-2) yields the solution

$$h(t, \tau) = \sum_{p,q} h_{pq} (t-\tau) e^{j2\pi(p\tau - q\tau)/T} \quad \forall t, \tau \in (-\infty, \infty) \quad (3-93)$$

to the integral equation

$$\int_{-\infty}^{\infty} k(t, s) h(s, \tau) ds + \lambda h(t, \tau) = g(t, \tau) \quad \forall t, \tau \in (-\infty, \infty) \quad (3-94)$$

where the $\{h_{pq}\}$ are the inverse Fourier transforms of the elements of the matrix

$$\underline{H}(f) = \begin{cases} [\underline{K}(f) + \lambda \underline{I}]^{-1} \underline{G}(f), & \forall f \in [-1/2T, 1/2T] \\ \underline{0}, & \forall f \notin [-1/2T, 1/2T] \end{cases} \quad (3-95)$$

where \underline{K} , \underline{G} are the Fourier transforms of the bandlimited (to $[-1/2T, 1/2T]$) matrices k , g which represent the T -periodic kernels

$$\begin{aligned} k(t,s) &= \sum_{p,q} k_{pq}(t-s) e^{j2\pi(pt-qs)/T} \\ g(t,s) &= \sum_{p,q} g_{pq}(t-s) e^{j2\pi(pt-qs)/T} \end{aligned} \quad (3-96)$$

as in Eqs. (3-83, 85).

Furthermore, if this solution is to be implemented as the impulse response of a linear system--as discussed in Section 2g--then the straightforward implementation shown in Fig. (3-12) suffices. Note that this linear system is composed of a matrix of time-invariant filters (with transfer functions $H_{pq}(f)$) which is flanked at the input and output, respectively, by the resolution and reconstruction systems of Fig. (3-11). Note also that since $\underline{H}(f)$ is bandlimited, then it can be exactly realized with a matrix sampled-data filter, with transfer function [53,54]

$$\underline{B}(f) = \sum_{p=-\infty}^{\infty} \underline{H}(f-p/T),$$

flanked at the input and output, respectively, by T -periodic impulse-samplers and ideal low-pass filters ($W(f)$).

Finally, it should be mentioned that such a system is not causal and cannot be physically implemented without the incorporation of a delay as discussed in Section 2g.

As an example of this integral-equation solution method, consider the special case (which occurs frequently in practice) where the kernels g and k in Eq. (3-94) are equal. Now if k is the autocorrelation function for the FDM signal discussed in the previous subsection (or has the same form), then the identical matrices \underline{G} , \underline{K} have the sparse form:

$$G_{pq}(f) = K_{pq}(f) = \begin{cases} \frac{1}{4} K_r(f), & |p| = |q| = N + r \\ 0, & \text{otherwise} \end{cases}$$

so that the solution matrix of Eq. (3-95) has elements given by the formula

$$H_{pq}(f) = \begin{cases} \frac{\frac{1}{4} K_r(f)}{\frac{1}{2} K_r(f) + \lambda}, & |p| = |q| = N + r \\ 0, & \text{otherwise.} \end{cases} \quad (3-97)$$

In the next chapter we will extend the integral-equation solution method of Eqs. (3-93)-(3-96) to include the case where the kernels k , g are not only jointly T -periodic in both variables, but also individually T -periodic in each variable (as is the case for the autocorrelation function of a random T -periodic signal), and where the interval of integration is finite.

4. Fourier Series Representation for Autocorrelation Functions

a) Definition. In this section we briefly consider the harmonic Fourier series representation for autocorrelation functions for cyclostationary processes. In contrast to the autocorrelation function-representations presented in the previous two sections, this representation does not result from a representation for cyclostationary processes: there is, in fact, no process-representation from which the Fourier series representation for an autocorrelation function automatically results.²²

If x is a cyclostationary process with period T , then its autocorrelation function is--by definition--jointly T -periodic in its two arguments:

$$k_{xx}(t+T, s+T) = k_{xx}(t, s) \quad \forall t, s.$$

Now, if we define a new function of two variables by making the change of variables:

$$r_{xx}(t', s') \triangleq k_{xx}(s' + t'/2, s' - t'/2),$$

then we have

$$k_{xx}(t, s) = r_{xx}(t-s, (t+s)/2) \quad (3-99)$$

so that the first argument in r_{xx} is the time-difference or size (and sense) of the time-interval (t, s) and the second argument is the time-average or midpoint of the time-interval (t, s) . Furthermore, the section

²²This is a conjecture which appears to be difficult to prove.

function $r_{xx}(t', \cdot)$ is T -periodic for every t' , and the section function $r_{xx}(\cdot, s')$ is a "decaying" function and will be contained in $L^2(-\infty, \infty)$ for most processes of interest; i.e., those processes whose correlation goes to zero faster than $1/|t-s|$.

We now represent the T -periodic functions $\{r_{xx}(t', \cdot); t' \in (-\infty, \infty)\}$ with their harmonic Fourier series:

$$r_{xx}(t', s') = \sum_{n=-\infty}^{\infty} c_n(t') e^{j2\pi n s' / T}$$

where

$$c_n(t') = \frac{1}{T} \int_{-T/2}^{T/2} r_{xx}(t', s') e^{-j2\pi n s' / T} ds' \quad (3-100)$$

If we substitute Eq. (3-100) into Eq. (3-99), then we obtain

$$k_{xx}(t, s) = \sum_{n=-\infty}^{\infty} c_n(t-s) e^{j\pi n(t+s)/T}$$

$$c_n(t) = \frac{1}{T} \int_{-T/2}^{T/2} k_{xx}(s+t/2, s-t/2) e^{-j2\pi n s / T} ds \quad (3-101)$$

which is the Fourier series representation (FSR) for the autocorrelation function.²³

Corresponding to this FSR for the autocorrelation function k_{xx} , we have the following representation for the double Fourier transform of K_{xx} :

23 This representation has been considered in great depth by Hurd [5]. In the third Chapter (which is devoted entirely to this representation) of his doctoral dissertation on cyclostationary processes, he investigates various modes of convergence of the partial sums of this Fourier series. However, neither Hurd, nor any of the previous investigators referenced by Hurd, appear to have identified the direct relationship existing between the Fourier-coefficient functions and the process as described in the remainder of this subsection.

$$K_{xx}(f, \nu) = \sum_n C_n(f-n/2T) \delta(\nu-f+n/T) \quad (3-102)$$

where C_n is the single Fourier transform of c_n . Notice that if we substitute this FSR into Eq. (3-85), then we obtain the following relationship between the elements of the correlation matrix for the HSR of Section 3, and the coefficient-functions for the FSR:

$$K_{pq}(f) = C_{p-q}(f+(p+q)/2T)W(f). \quad (3-103)$$

Furthermore, by equating the right members of Eqs. (3-102), (3-84), we obtain the relation:

$$C_n(f-n/2T) = \sum_p K_{(p-n)p}(f-p/T), \quad (3-104)$$

so that the n^{th} FSR coefficient function is equal to the sum of the frequency-translated versions of the elements of the n^{th} off-diagonal in the HSR correlation matrix $K(f)$.

In view of the fact that the HSR and the FSR for the autocorrelation functions of CS processes both employ the harmonic exponentials as basis functions, one might expect to find some interesting relationships between the two representations--in addition to Eqs. (3-103), (3-104). The following is perhaps the most fundamental relationship existing between the HSR and the FSR:

From Section 3, we know that the elements of the correlation matrix for the HSR are--by definition--cross-correlation functions of jointly WSS complex processes:

$$k_{pq}(t-s) = E\{a_p(t)a_q^*(s)\}$$

where the process $a_p(t)$ is obtained by reducing the T-CS process $x(t)e^{-j2\pi pt/T}$ to a WSS process via bandlimiting to the frequency-interval $[-1/2T, 1/2T]$ (Theorem (3-7)). Similarly, the FSR coefficient-functions are crosscorrelation functions of WSS complex processes:

$$c_p(t-s) = E\{b_p(t)b_p(s)\}$$

where the process $b_p(t)$ is obtained by reducing the T-CS process $x(t)e^{-j\pi pt/T}$ to a WSS process via phase-randomization.

We state this new result as a theorem:

THEOREM(3-8): Let x be a T-CS process whose autocorrelation function admits the FSR

$$\begin{aligned} k_{xx}(t,s) &= \sum_{p=-\infty}^{\infty} c_p(t-s)e^{j\pi p(t+s)/T} \\ c_p(t) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{xx}(s+t/2, s-t/2)e^{-j2\pi ps/T} ds, \end{aligned} \quad (3-105)$$

and define b_p to be the WSS complex process obtained by phase-randomizing the T-CS complex process x_p :

$$b_p(t) \triangleq x_p(t+\theta)$$

where

$$x_p(t) = x(t)e^{-j\pi pt/T}$$

and where the PDF for θ is uniform:

$$p_{\theta}(\sigma) = \begin{cases} 1/T, & |\sigma| \leq T/2 \\ 0, & |\sigma| > T/2 \end{cases}.$$

Now, the coefficient functions in the FSR are equal to the cross-correlation functions of the $\{b_p\}$:

$$c_p(t-s) = E\{b_p(t)b_p(s)\}. \quad (3-106)$$

Proof:

$$\begin{aligned} E\{b_p(t)b_p(s)\} &= E\{x_p(t+\theta)x_p(s+\theta)\} \\ &= \int_{-\infty}^{\infty} E\{x_p(t+\sigma)x_p(s+\sigma)\}P_{\theta}(\sigma)d\sigma \\ &= \frac{1}{T} \int_{-T/2}^{T/2} E\{x(t+\sigma)x(s+\sigma)\}e^{-j\pi p(t+s+2\sigma)/T}d\sigma \\ &= \frac{1}{T} \int_{-T/2}^{T/2} k_{xx}(t+\sigma,s+\sigma)e^{-j\pi p(t+s)/T}e^{-j2\pi p\sigma/T}d\sigma \\ &= \frac{1}{T} \int_{-T/2}^{T/2} \sum_q c_q(t-s)e^{j\pi(q-p)(t+s)/T}e^{j2\pi(q-p)\sigma/T}d\sigma \\ &= \sum_q c_q(t-s)e^{j\pi(q-p)(t+s)/T} \frac{1}{T} \int_{-T/2}^{T/2} e^{j2\pi(q-p)\sigma/T}d\sigma \\ &= c_p(t-s). \end{aligned}$$

QED

Note that since b_p is a complex process (for $p \neq 0$), then $c_p(t-s)$ is not the autocorrelation function for b_p , but rather the crosscorrelation function for b_p and b_p^* . Note also that since x is real, then $b_p^* = b_{-p}$.

The system which transforms any T-CS process x into the WSS processes $\{b_p\}$ is shown in Fig. (3-13) where θ is a delay line whose delay length is a random variable uniformly distributed over the interval $[-T/2, T/2]$. Hence, unlike the resolution systems for the TSR's and the HSR, this

transformation is a random periodically time-varying linear system (as discussed in Section 5 of Chapter II), and has random impulse-response function

$$g_p(t, \tau) = e^{-j\pi p \tau / T} \delta(t + \theta - \tau).$$

Observe that this random system has the same form as the HSR resolution system of Fig. (3-11) except that "phase-randomizers" are used in place of "bandlimiters".

The p^{th} inverse transformation which converts b_p back into the original process x is also a random periodically time-varying linear system, and is statistically dependent on the resolution transformation as well as on its input b_p . The random impulse-response function for this inverse transformation is

$$g_p^{-1}(t, \tau) = e^{j\pi p \tau / T} \delta(t - \theta - \tau),$$

so that the output process is

$$x(t) = b_p(t - \theta) e^{j\pi p t / T}$$

where θ is the same random-phase variable employed (in g_p) to resolve x into b_p .

b) Properties of the Fourier series representation. In this subsection, we present various properties of the FSR which are useful in analyses dealing with CS processes:

i) Phase randomization and stationarity. If the process \tilde{x} is obtained from the T-CS process x by the introduction of a random-phase variable θ :

$$\tilde{x}(t) = x(t+\theta)$$

then the FSR coefficient-functions $\{\tilde{c}_n\}$ for the autocorrelation function for \tilde{x} are given by

$$\tilde{c}_n(t) = P_\theta(n/T) c_n(t) \quad (3-107)$$

where $\{c_n\}$ are the FSR coefficient-functions for x , and $P_\theta(\cdot)$ is the Fourier transform of the PDF for θ , $p_\theta(\cdot)$. Now, if the PDF is uniform, then

$$P_\theta(n/T) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

so that

$$\tilde{c}_n(t) = \begin{cases} c_0(t), & n = 0 \\ 0, & n \neq 0 \end{cases}$$

and

$$k_{\tilde{x}\tilde{x}}(t-s) = c_0(t-s).$$

Hence, the zeroth coefficient in the FSR is the autocorrelation function for the WSS process obtained via phase-randomization. This result is more directly obtained by noting that

$$\begin{aligned} c_0(t-s) &= E\{b_0(t)b_0(s)\} \\ &= E\{x(t+\theta)x(s+\theta)\} \end{aligned} \quad (3-108)$$

where θ is the uniformly distributed random-phase variable.

ii) Filtering and bandlimiting. If \tilde{x} is the output of a linear time-invariant filter with transfer function $G(f)$, and with T-CS input x , then the Fourier transform of the FSR coefficient-functions for \tilde{x} are given by:

$$\tilde{C}_n(f) = G(f+n/2T)G^*(f-n/2T)C_n(f) \quad (3-109)$$

where $\{C_n\}$ are the Fourier transforms of the FSR coefficient-functions for x . Now, if G is a bandlimiting filter with bandwidth $2B$:

$$G(f) = 0, \quad |f| > B,$$

then

$$\tilde{C}_n(f) \equiv 0, \quad n \geq 2BT,$$

so that bandlimited processes possess only a finite number of non-zero coefficient-functions. This is, of course, true for "band-pass" as well as "low-pass" processes. Note that if $B \leq 1/2T$, then only the zeroth coefficient-function is non-zero and the bandlimited process is WSS--a result previously obtained.

iii) Constraints on coefficient-functions. There are a number of interesting properties possessed by the FSR coefficient-functions. These properties can be viewed as constraints which must be satisfied by the coefficient-functions corresponding to any CS process. We briefly discuss three particularly interesting types of constraints:

(1) Symmetry constraints: Using the defining equation (Eq. (3-102)), for the FSR, one can easily show that the coefficient-functions $\{c_n\}$ and their Fourier transforms $\{C_n\}$ exhibit the symmetries:

$$\begin{aligned} c_n(t) &= c_{-n}^*(-t) = c_n(-t) \\ C_n(f) &= C_{-n}^*(f) = C_n(-f) \end{aligned} \quad (3-110)$$

where the last equality in each line is valid only if x is a real process.

(2) Norm constraints: Employing the Cauchy-Schwarz[2] inequality for inner products (in H_{RV}) results in

$$k_{xx}^2(t,s) \leq k_{xx}(t,t)k_{xx}(s,s).$$

Now, using the FSR for k_{xx} and replacing t by $(\sigma+t/2)$, and s by $(\sigma-t/2)$, and integrating w.r.t. σ over the interval $[-T/2, T/2]$ yields the bound

$$\sum_n |c_n(t)|^2 \leq \sum_n |c_n(0)|^2 e^{j2\pi nt/T} \leq \sum_n |c_n(0)|^2. \quad (3-111)$$

Furthermore, if we again integrate (over $[-NT/2, NT/2]$), then we obtain the bound

$$\sum_n \frac{1}{NT} \int_{-NT/2}^{NT/2} |c_n(t)|^2 dt \leq |c_0(0)|^2. \quad (3-112)$$

(3) Support constraints: Since the mean-squared value of any random variable must be non-negative, then by considering the particular random variable

$$y = \int_{-\infty}^{\infty} g(t)x(t)dt,$$

where x is a T-CS process, and $g(\cdot)$ is any real function, we obtain the non-negativity condition

$$E\{y^2\} = \iint_{-\infty}^{\infty} g(t)g(\tau)k_{xx}(t,\tau)dtd\tau \geq 0 \quad (3-113)$$

Now, using Parseval's relation [2], we obtain

$$E\{y^2\} = \iint_{-\infty}^{\infty} G(f)G^*(v)K_{xx}(f,v)dfdv \geq 0 \quad (3-114)$$

and, using the FSR of Eq. (3-113) for K_{xx} yields

$$\sum_n \int_{-\infty}^{\infty} G(f+n/2T)G^*(f-n/2T)C_n(f)df \geq 0. \quad (3-115)$$

This general non-negativity condition imposes various interesting constraints on the support of the Fourier transforms of the FSR coefficient-functions. Specifically, if we choose a "double band-pass" shape for G :

$$G(f) = \begin{cases} A, & f \in [m/2T - f_0 - B, m/2T - f_0 + B] \cup [m/2T + f_0 - B, m/2T + f_0 + B] \\ A^*, & f \in [-m/2T - f_0 - B, -m/2T - f_0 + B] \cup [-m/2T + f_0 - B, -m/2T + f_0 + B] \\ 0, & \text{otherwise} \end{cases}$$

where

$$0 \leq f_0 + B \leq 1/4T,$$

then, Eq. (3-124) becomes

$$\begin{aligned} \operatorname{Re}\{A^2 \int_{-f_0-B}^{-f_0+B} C_m(f) df + A^2 \int_{f_0-B}^{f_0+B} C_m(f) df\} + |A|^2 \int_{-f_0-B}^{-f_0+B} C_0(f+m/2T) df \\ + |A|^2 \int_{f_0-B}^{f_0+B} C_0(f+m/2T) df \geq 0. \end{aligned}$$

Now, by choosing A equal to 1 and to j , we obtain the constraint

$$\int_{-B}^B [C_0(f+m/2T+f_0) + C_0(f+m/2T-f_0)] df \geq \left| \int_{-B}^B \operatorname{Re}\{C_m(f+f_0) + C_m(f-f_0)\} df \right|$$

where,

$$0 \leq f_0 + B \leq 1/4T. \quad (3-116)$$

Among other things, this is a constraint on the support of the FSR coefficient-functions. From it, we conclude that if the m^{th} coefficient-function C_m has support throughout the frequency band $[-1/4T, 1/4T]$, then the zeroth coefficient-function C_0 must have support throughout the frequency band $[m/2T - 1/4T, m/2T + 1/4T]$. Other constraints can be derived by choosing other types of gate-functions for G . In fact, it appears as though the m^{th} coefficient-function C_m will have support throughout the frequency band $[-B, B]$ for any B if and only if the zeroth coefficient-function has support throughout the band $[m/2T - B, m/2T + B]$.

c) Solution of integral equations, and realization of periodically time-varying systems. In Subsections 2f, 3e, we saw that series representations for autocorrelation functions (and other periodic functions of two variables) can be employed for solving linear integral equations in which these functions appear. This was possible for the TSR's because of the separation of variables effected by the representation. Although the FSR does not separate variables, it does provide a means for converting the integral equation to a set of linear algebraic equations. However, as we will see, finite order FSR's do not convert integral equations to a finite set of algebraic equations as was the case for TSR's. The set of algebraic equations obtained with the FSR is infinite, and not particularly easy to solve except in special cases.

Using Parseval's relation [2], the linear integral equation

$$\int_{-\infty}^{\infty} h(t, s) k(s, \tau) ds + \lambda h(t, \tau) = g(t, \tau) \quad \forall t, \tau \in (-\infty, \infty) \quad (3-117)$$

can be converted to the equivalent (dual) linear integral equation

$$\int_{-\infty}^{\infty} H(f, \nu) K(\nu, \omega) d\nu + \lambda H(f, \omega) = G(f, \omega), \quad \forall f, \omega \in (-\infty, \infty) \quad (3-118)$$

where K , H , G are the double Fourier transforms of k , h , g . Now, if k, g are jointly T -periodic in their two variables, as are the autocorrelation and crosscorrelation functions for T -CS processes, then they admit FSR's, and their double Fourier transforms can be expressed in the form

$$\begin{aligned} K(f, \nu) &= \sum_n C_n(f-n/2T) \delta(\nu-f+n/T) \\ G(f, \nu) &= \sum_n G_n(f-n/2T) \delta(\nu-f+n/T). \end{aligned} \quad (3-119)$$

Assuming the same form

$$H(f, \nu) = \sum_n H_n(f-n/2T) \delta(\nu-f+n/T) \quad (3-120)$$

for the solution function, and substituting into the integral equation yields, after integration, the following equivalent set of linear algebraic equations:

$$\begin{aligned} \sum_{m=-\infty}^{\infty} C_{n-m}(f-(n+m)/2T) H_m(f-m/2T) + \lambda H_n(f-n/2T) \\ = G_n(f-n/2T) \quad \forall n. \end{aligned} \quad (3-121)$$

There does not appear to be any way to obtain a general closed form solution to this infinite set of equations;²⁴ however, if $\{C_n\}$, $\{G_n\}$ satisfy certain bandlimiting constraints, then solutions are possible. We consider the special (but common) case where $g = k$. Now, if the $\{C_n\}$ satisfy the constraint:

24 This difficulty appears to be related to the fact (conjecture) that the FSR for autocorrelation functions does not result from a process-representation as do the HSR and TSR's.

$$C_n(f) = 0, \quad |f| \geq 1/4T, \quad n \neq 0 \quad (3-122)$$

then, Eqs. (3-121) admit the solutions:

$$H_n(f-n/2T) = \frac{C_n(f-n/2T)[1-H_0(f)]}{\lambda + C_0(f-n/T)}, \quad n \neq 0$$

$$H_0(f) = \frac{C_0(f) - S(f)}{\lambda + C_0(f) - S(f)},$$

where

$$S(f) \triangleq \sum_{m \neq 0} \frac{|C_m(f-m/2T)|^2}{\lambda + C_0(f-m/T)}, \quad (3-123)$$

and the solution to the original integral equation is

$$h(t, \tau) = \sum_n h_n(t-\tau) e^{j\pi n(t+\tau)/T}$$

where $\{h_n\}$ are the inverse Fourier transforms of the $\{H_n\}$.

If the bandlimiting constraint, Eq. (3-122), is weakened to

$$C_n(f) = 0, \quad |f| \geq M/4T, \quad n \neq 0,$$

then the solutions to Eq. (3-121) become:

$$H_n(f-n/2T) = \frac{C_n(f-n/2T)[1-H_0(f)]}{\lambda + C_0(f-n/T)}, \quad n \neq 0, \pm 1, \dots, \pm(M-1)$$

and $\{H_n; n = 0, \pm 1, \dots, \pm(M-1)\}$ can be obtained by inverting a $(2M-1) \times (2M-1)$ matrix of functions.

As a simple example of this integral-equation solution method, consider the autocorrelation function corresponding to an AM process (see Sec. 2 of Chapter II) as a kernel:

$$\begin{aligned}
 k(t,s) &= \cos(2\pi t/T) \cos(2\pi s/T) k_{yy}(t-s) \\
 &= \sum_n c_n(t-s) e^{j\pi n(t+s)/T}
 \end{aligned}$$

where,

$$c_0(t) = \frac{1}{2} k_{yy}(t) \cos(2\pi t/T)$$

$$c_2(t) = c_{-2}(t) = \frac{1}{4} k_{yy}(t)$$

$$c_n(t) = 0, \quad n \neq 0, \pm 2.$$

The Fourier transforms of these FSR coefficient-functions are:

$$C_0(f) = \frac{1}{4} K_{yy}(f-1/T) + \frac{1}{4} K_{yy}(f+1/T)$$

$$C_2(f) = C_{-2}(f) = \frac{1}{4} K_{yy}(f)$$

$$C_n(f) = 0, \quad n \neq 0, \pm 2.$$

Now, if K_{yy} (the power spectral density of the "baseband" signal) is bandlimited to $[-1/4T, 1/4T]$, then the constraint of Eq. (3-122) is satisfied, and substituting into Eqs. (3-123) yields the solution coefficient-functions:

$$H_0(f) = \frac{1}{4} R(f-1/T) + \frac{1}{4} R(f+1/T)$$

$$H_2(f) = H_{-2}(f) = \frac{1}{4} R(f)$$

$$H_n(f) = 0, \quad n \neq 0, \pm 2$$

where

$$R(f) \triangleq \frac{2K_{yy}(f)}{2\lambda + K_{yy}(f)}.$$

Hence, the solution to the original integral equation is:

$$h(t, \tau) = \cos(2\pi t/T) \cos(2\pi \tau/T) r(t - \tau),$$

where $r(\cdot)$ is the inverse Fourier transform of $R(\cdot)$.

If the solution to our integral equation is to be implemented as the impulse response of a linear system--as discussed in Subsections 2g, 3e--then the straightforward FSR realization shown in Fig. (3-14a), can be employed. The solution to the integral equation will not, in general, be the impulse-response function of a causal system, so that the time-invariant path-filters shown in Fig. (3-14a) will not be causal. However, if one is given a causal impulse-response function h to start with, then the coefficient-functions $\{h_n\}$ in its FSR,

$$h_n(t) = \frac{1}{T} \int_{-T/2}^{T/2} h(s+t/2, s-t/2) e^{-j2\pi ns/T} ds, \quad (3-124)$$

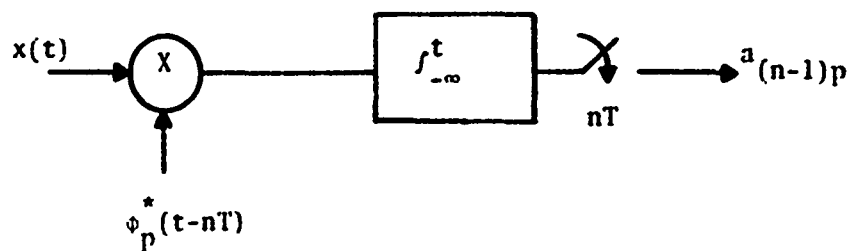
$$= 0 \text{ for } t < 0.$$

will be the impulse-response functions of causal time-invariant filters, so that the realization shown in Fig. (3-14a) will then be causal.

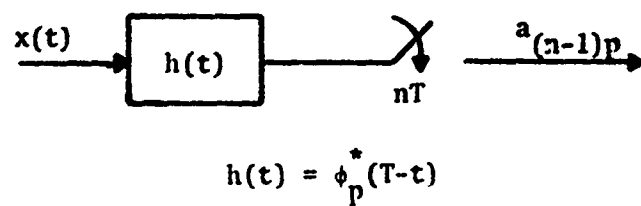
This FSR realization can be simplified for the purpose of implementation. Such a simplified version is shown in Fig. (3-14b) where only input modulators are used, and all devices are real. The impulse-response functions for the path-filters with transfer functions G_{n1} , G_{n2} are related to the original FSR coefficient-functions as follows:

$$g_{n1}(t) = 2\operatorname{Re}\{h_n(t)\}\cos\pi nt/T - 2\operatorname{Im}\{h_n(t)\}\sin\pi nt/T$$

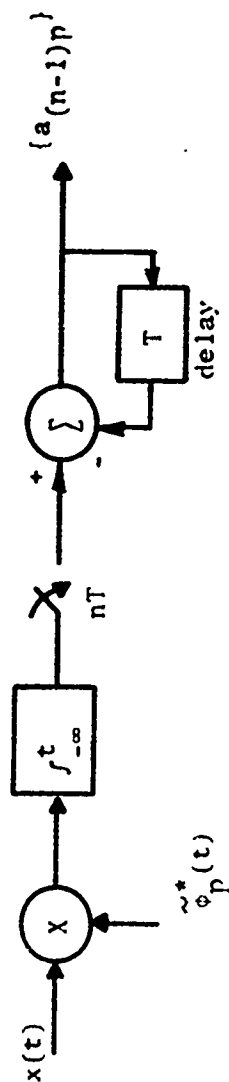
$$g_{n2}(t) = -2\operatorname{Re}\{h_n(t)\}\sin\pi nt/T - 2\operatorname{Im}\{h_n(t)\}\cos\pi nt/T \quad (3-125)$$



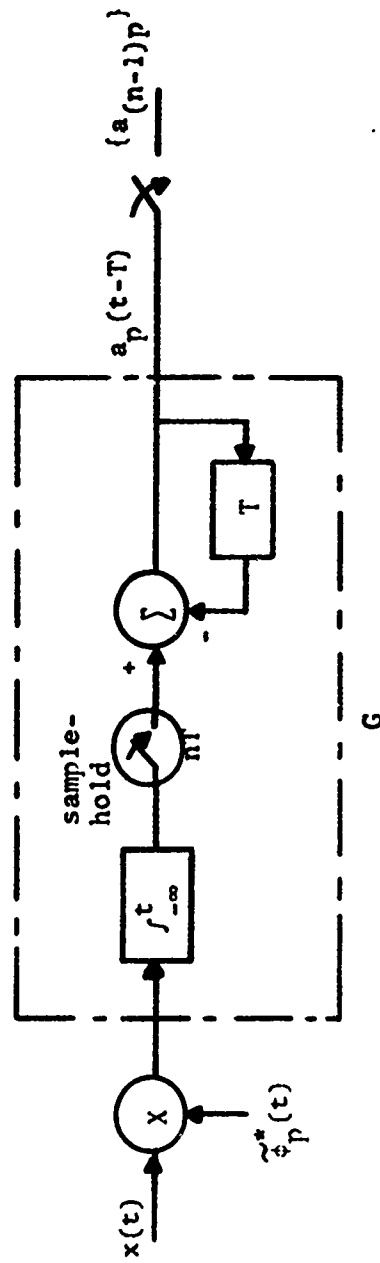
Figure(3-1) Multiplier-integrator realization of the process-resolution device.



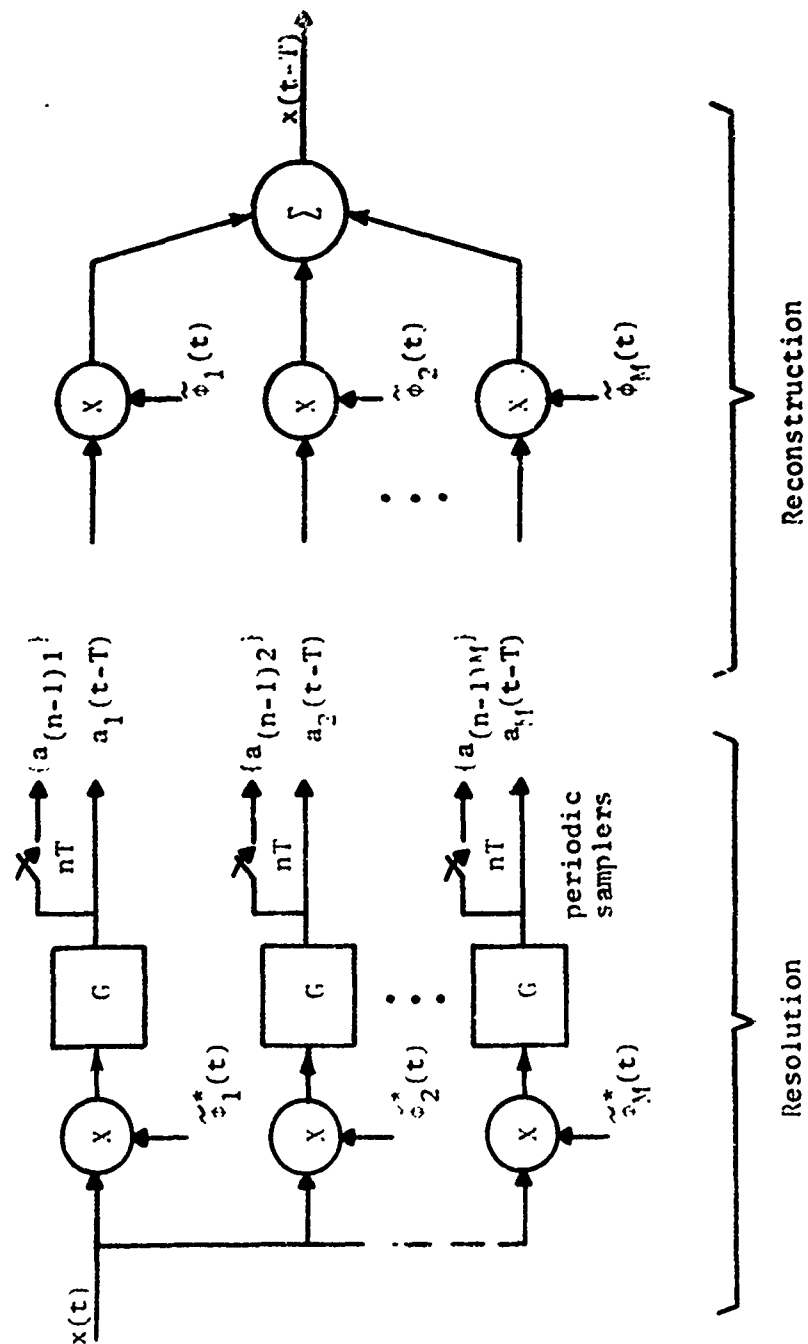
Figure(3-2) Filter-sample realization of the process-resolution device.



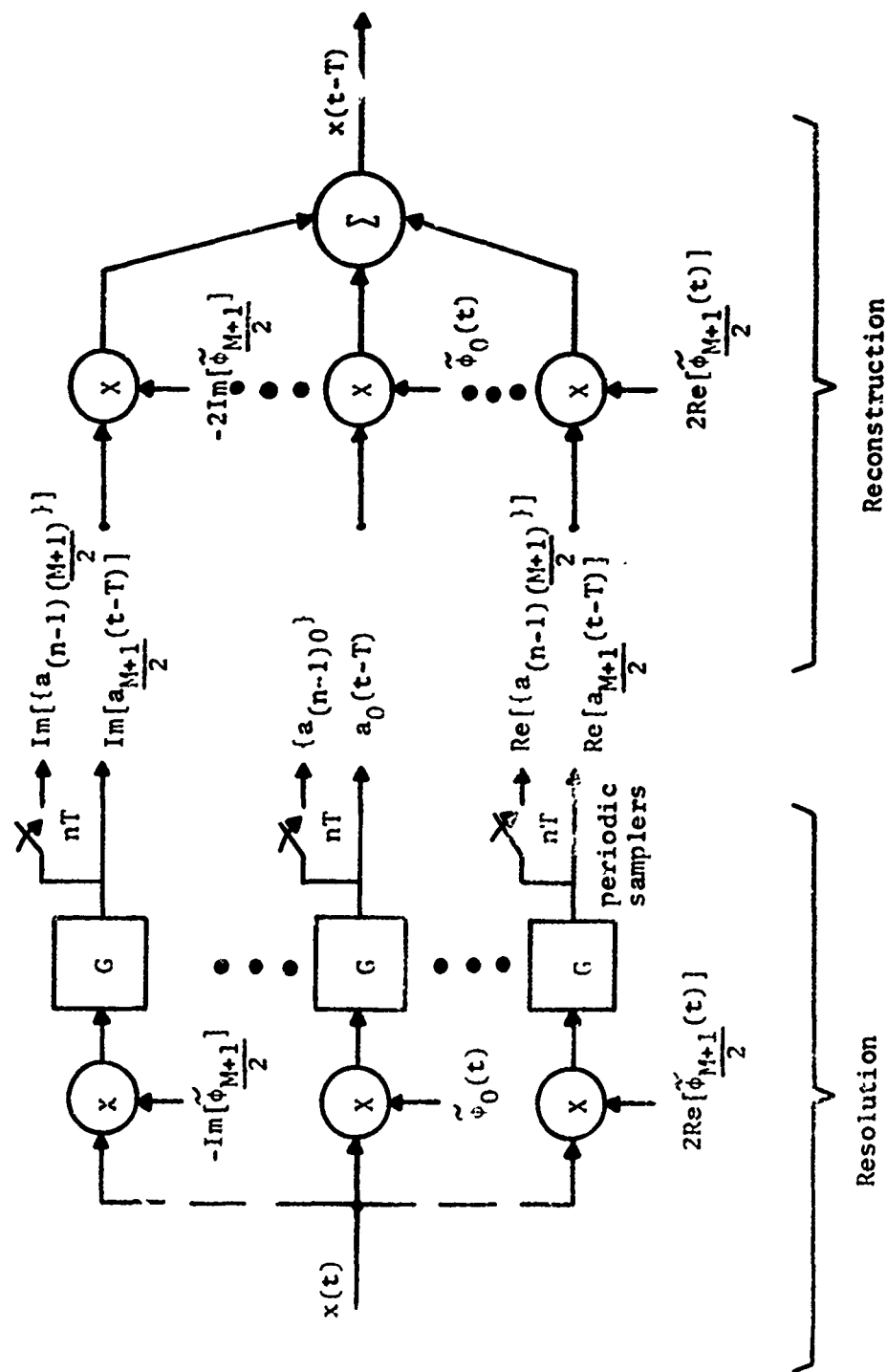
Figure(3-3) Adaptation of multiplier-integrator realization of the resolution device for continuous operation. $(\tilde{\phi}_p(t) = \sum_i \phi_p(t-iT))$



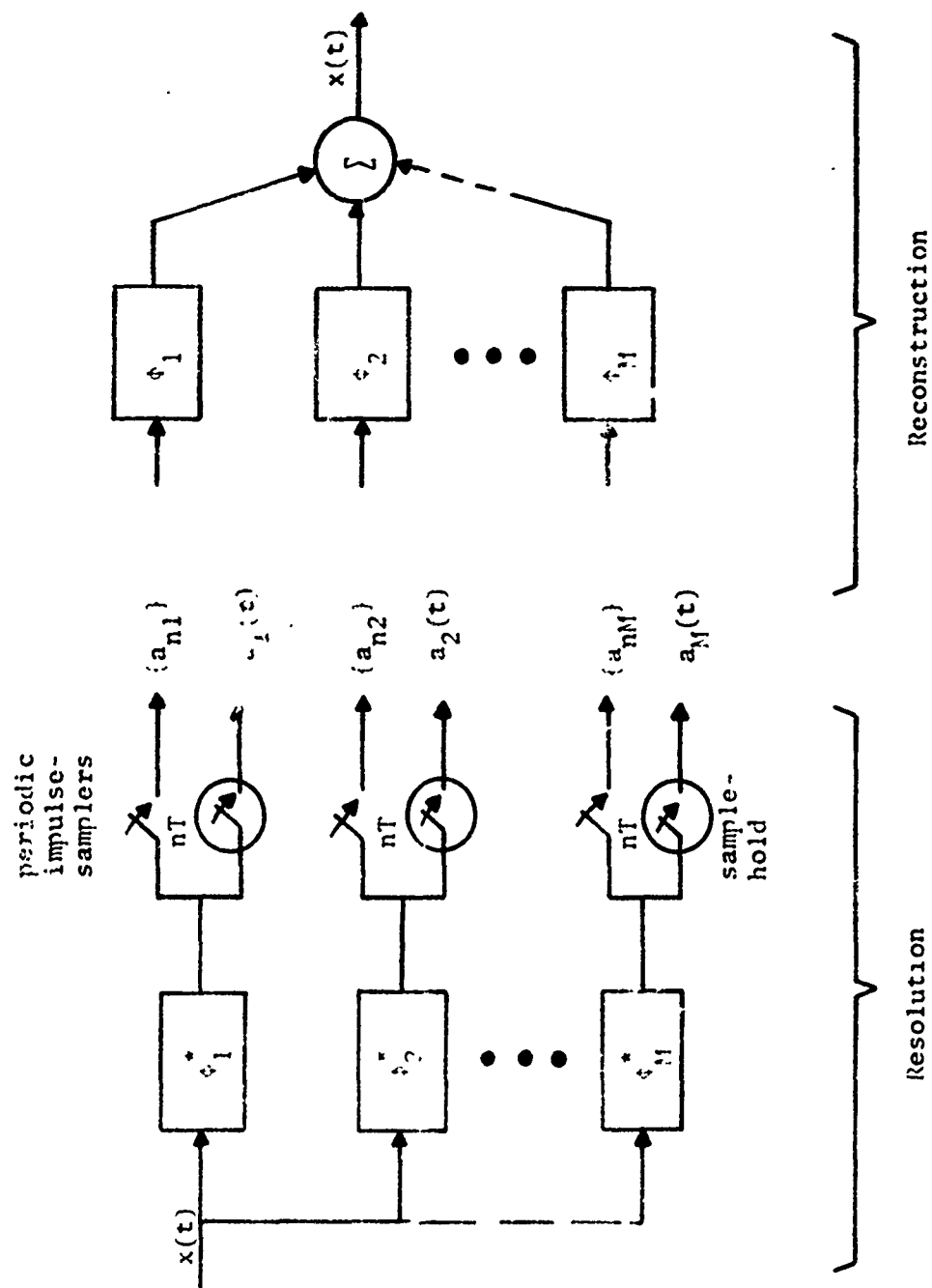
Figure(3-4) Realization of resolution device for computing a continuous (PAM) representor $a_p(t)$ and a discrete representor $\{a_{np}\}$.



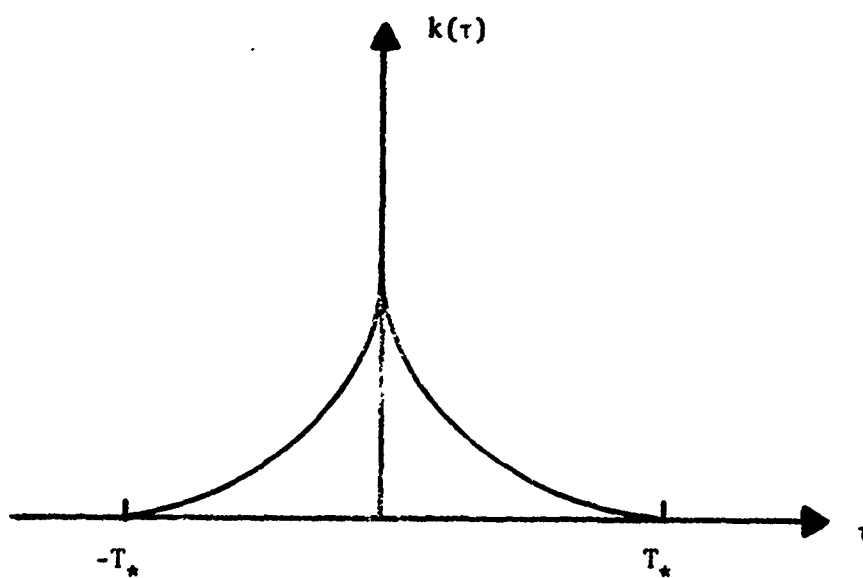
Figure(3-5) TSR resolution and reconstruction of a T-CS random process: $\{a_{np}\}$ are the discrete representors, and $\{a_p(t)\}$ are the continuous representors. (G is defined in Fig.(3-4))



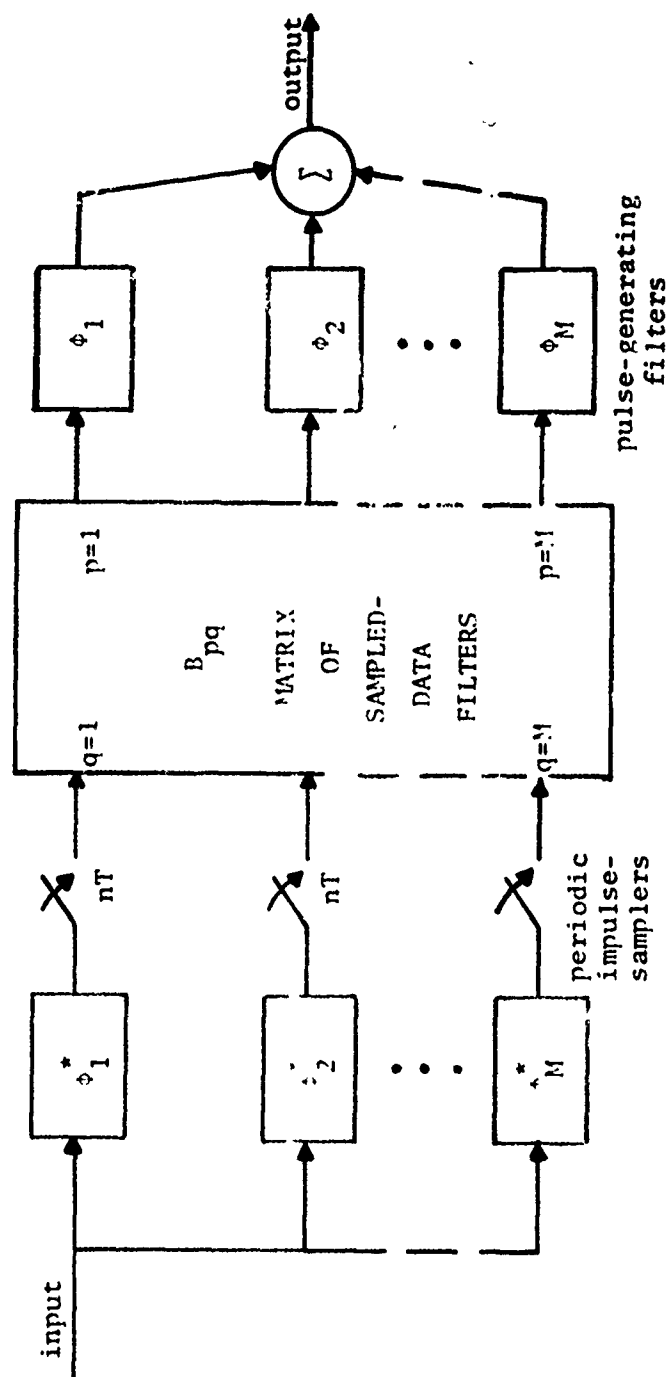
Figure(3-6) TSR resolution into real and imaginary parts of the representors of a T-CS process.



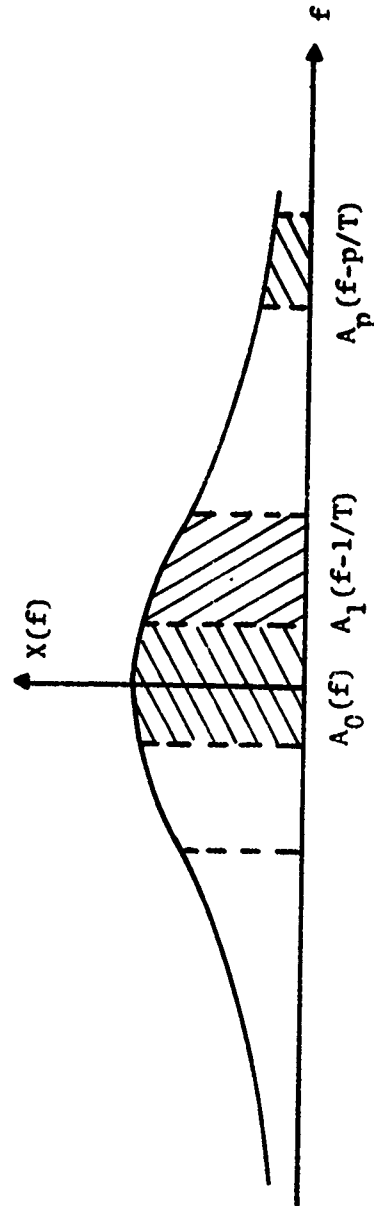
Figure(3-7) TSR resolution and reconstruction of a T-CS random process. (The $\{\phi_p\}$ are transfer functions of time-invariant filters.)



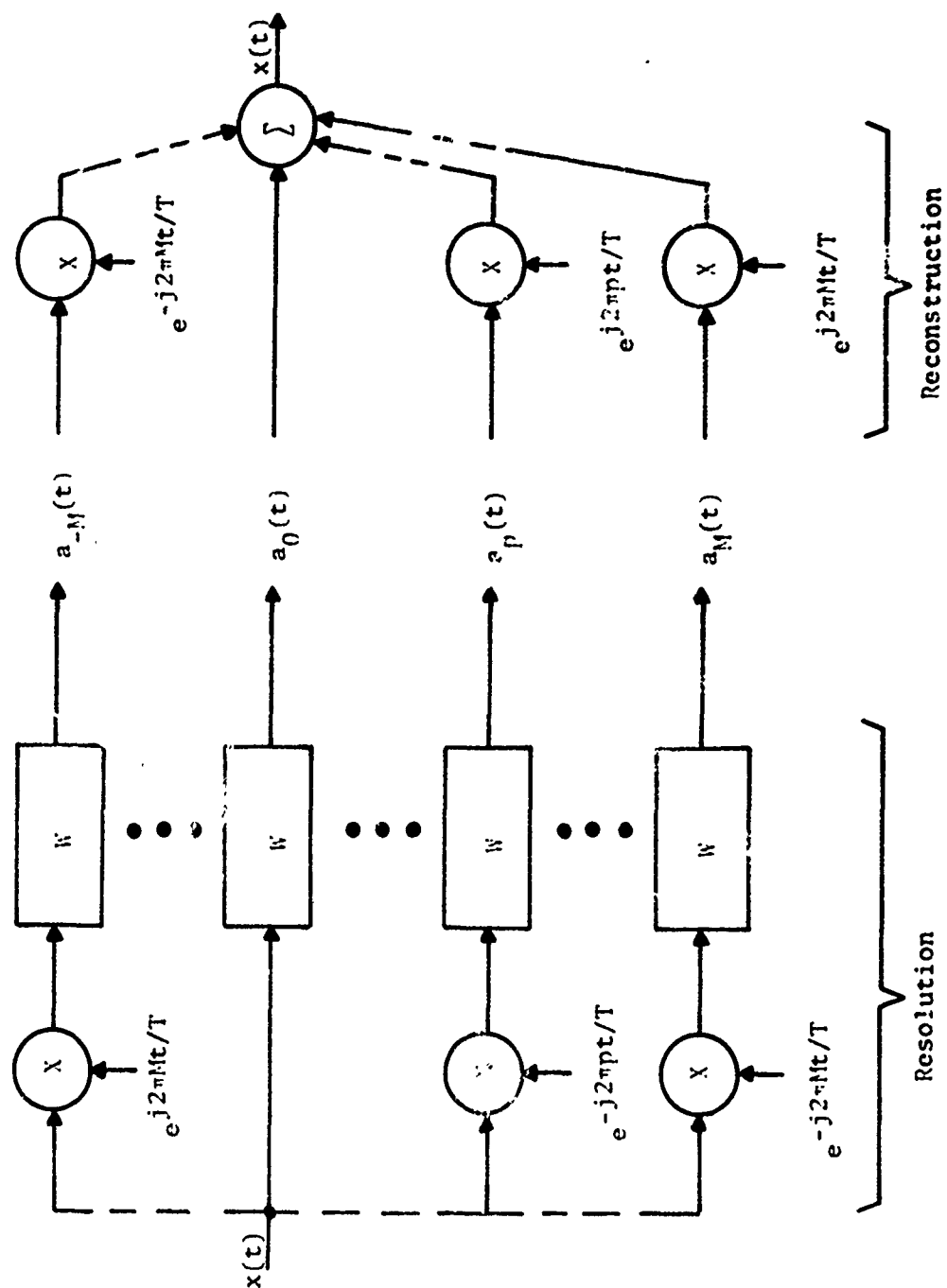
Figure(3-8) Autocorrelation function shape as described in Section 2d (for $n > 1$).



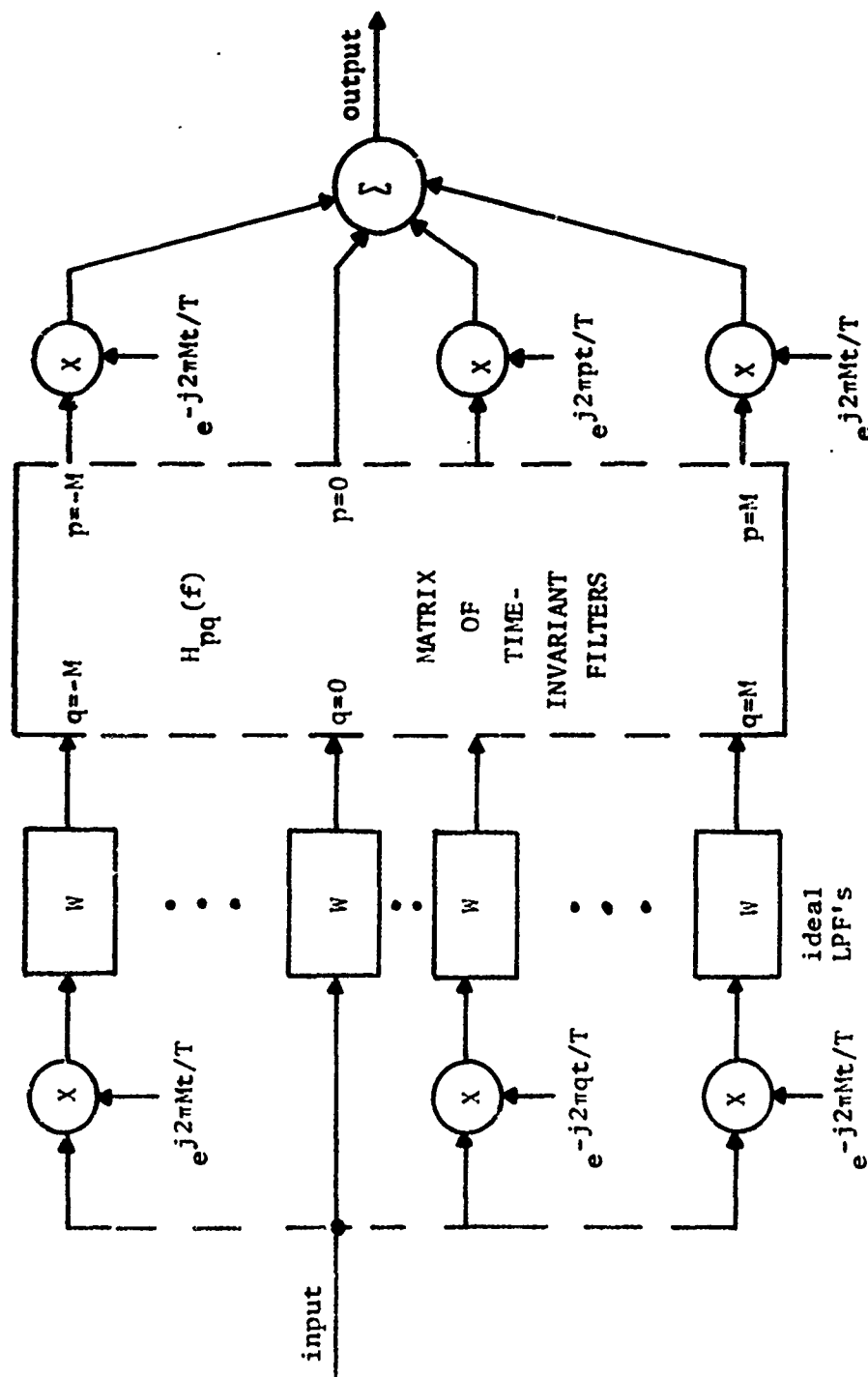
Figure(5-9) Implementation of a periodically(T) time-varying linear system whose impulse-response function has an N^{th} order translation series representation.



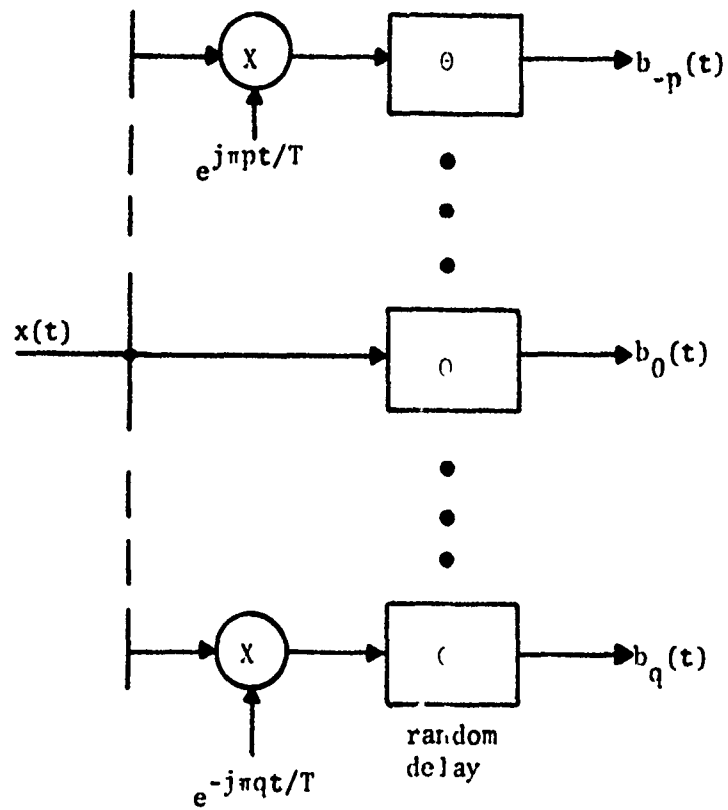
Figure(3-10) Resolution of a hypothetical Fourier transform of a T-CS process into its HSR representors.



Figure(3-11) HSR resolution and reconstruction of a T-CS process. (W is an ideal low-pass filter with cutoff frequency $1/2T$.)



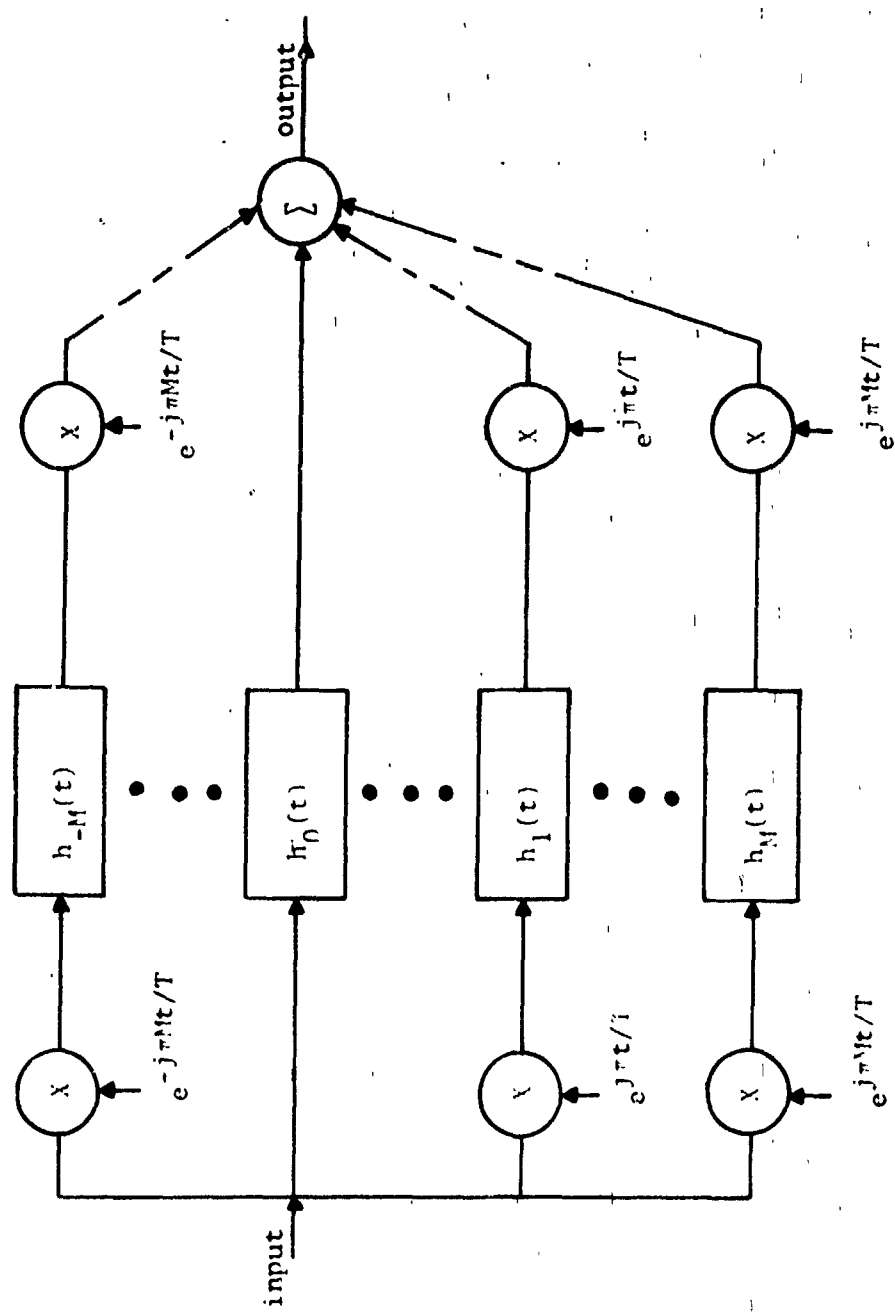
Figure(3-12) Implementation of a periodically(T) time-varying linear system whose impulse-response function has an M^{th} order harmonic series representation.



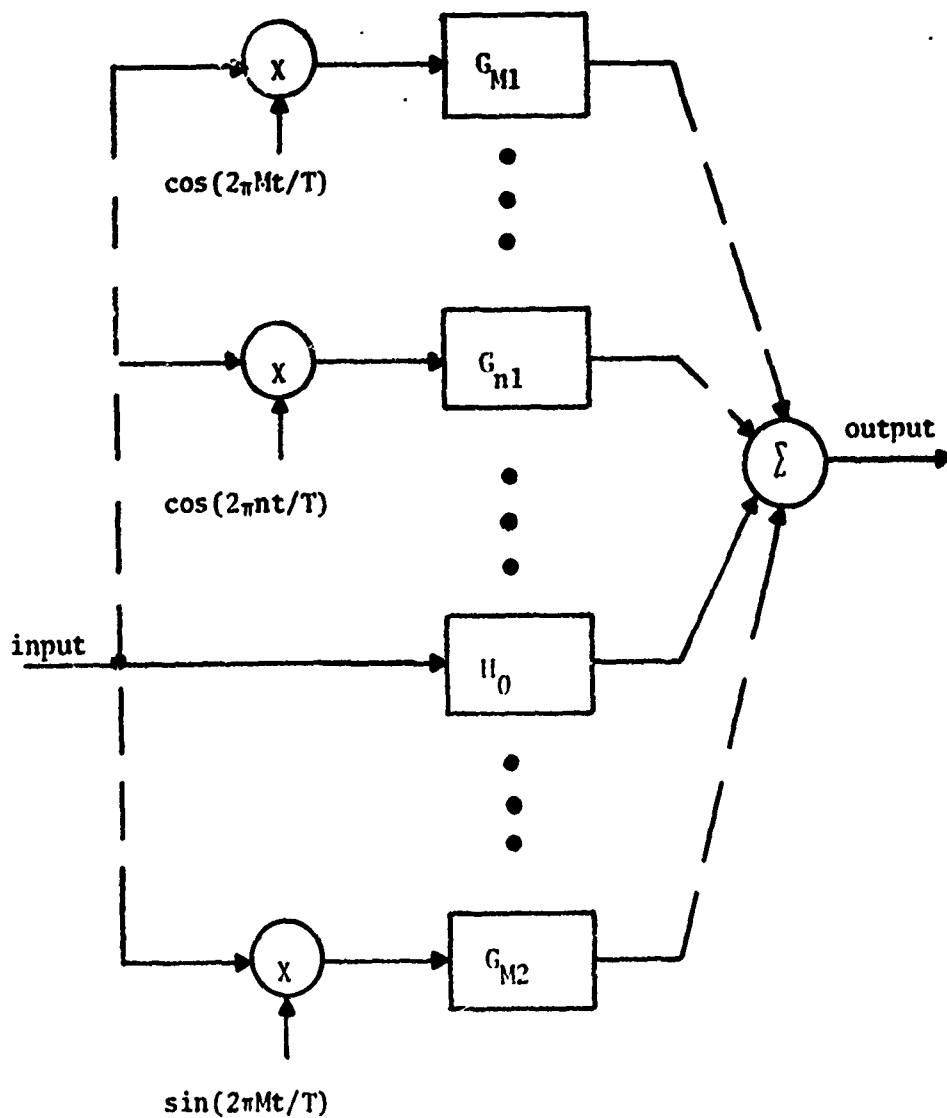
$$k_{xx}(t,s) = \sum_p c_p(t-s) e^{j\pi p(t+s)/T}$$

$$c_p(t-s) = E\{b_p(t)b_p(s)\}$$

Figure(3-13) Transformation of WSS process x into WSS complex processes $\{b_p\}$ whose crosscorrelations are the FSR coefficient-functions $\{c_p\}$. (The random delay θ is uniformly distributed over $[0,T]$.)



Figure(3-14a) Realization of a periodically(T) time-varying linear system whose impulse-response function has an M^{th} order Fourier series representation.-



Figure(3-14b) Simplified realization of a periodically(T) time-varying linear system whose impulse-response function has an M^{th} order FSR.

CHAPTER IV

LEAST-MEAN-SQUARED-ERROR LINEAR ESTIMATION OF CYCLOSTATIONARY PROCESSES

1. Introduction

The material in the preceding two chapters serves mainly as a theoretical base upon which various problems dealing with random periodic phenomena can be formulated and solved. Chapter II provides a base for obtaining mathematical models for physical systems and signals, so that physical problems can be formulated mathematically, and Chapter III provides a base for simplifying the mathematical models so that analyses can be simplified, and solutions obtained.

In this fourth chapter, we consider a particular application of the theory developed in these earlier chapters. Although this application is but one of many possibilities, it has special significance in that it has been the primary motivation for developing the theory. Furthermore, the problem considered in this application is of significant practical (as well as theoretical) value in that its solution provides a means for improving the quality of present communication systems, and a means for evaluating certain performance measures for newly proposed (as well as existing) communication systems, and it leads to useful insight into the nature of optimum signal processing.

Specifically, the problem to be considered is that of least-mean-squared-error linear estimation of imperfectly observed cyclostationary processes--or, in communications jargon, optimum filtering of noisy

distorted cyclostationary signals. The general form of this optimum filtering problem is shown pictorially in Fig. (4-1): A random signal process x is transmitted through (and distorted by) a channel which is composed of a dispersive transformation which can be nonlinear, time-varying, or even random; and noise which can be additive and/or multiplicative. The received signal y (the distorted version of x) is processed by an "optimum filter" which is designed to minimize the mean-squared difference between its output $\hat{x}(t)$ and the transmitted signal $x(t)$ for all values of t in some interval T_0 (usually $(-\infty, \infty)$).

Typically, the filter is chosen to be optimum subject to the constraint that it be linear. Nonlinear processing of the received signal is usually carried out in the receiver which follows the optimum filter. Such nonlinear signal processing will not be considered here. It should be mentioned, however, that if the transmitted signal x is a Gaussian process, and if the channel is composed of a linear deterministic transformation and additive Gaussian noise, then the optimum (least-mean-squared error) filter subject to no constraints will be a linear filter [51].

The particular aspects of the optimum-filtering problem that will be investigated in this chapter are the structure and the performance of optimum (linear) filters for cyclostationary signals. In fact, we will prove (and demonstrate with numerous examples) the proposition that optimum filters for cyclostationary signals are periodically time-varying systems which can yield significantly improved performance over optimum filters which are constrained to be time-invariant and which thereby

ignore the cyclic character of the cyclostationary signals.²⁵ We will also prove (and demonstrate with examples) the fact that optimum linear receivers consist of demultiplexor-demodulators followed by time-invariant filters--the order of these operations being the reverse of that sometimes employed in practice. These results are of considerable importance in view of the great preponderance of cyclostationary signals in communication systems (as emphasized throughout Chapter II). Specific examples which are worked out in detail include, among others, signals which are amplitude-modulated, frequency-shift-keyed, pulse-amplitude-modulated, frequency-division-multiplexed, and time-division-multiplexed.

Before proceeding, we briefly discuss the origin of the above proposition on improved performance: In Section 5 of Chapter II, it was shown that any cyclostationary process can be reduced to a stationary process by randomizing the phase (or time-origin) of the process (Theorem (2-15)), and that the resultant autocorrelation function is precisely the same as the function obtained by averaging, over one period, the original autocorrelation function of the cyclostationary process (Theorem (2-16)). This has long been known (as brought out in the historical notes of Section 3 of Chapter I) and indiscriminantly used to simplify analyses dealing with cyclostationary processes. In fact most systems-analysts, in solving for optimum filters for cyclostationary signals, use the stationarized model for their cyclostationary signals, and come up with optimum time-invariant (non-coherent)

²⁵ To my knowledge, the only other result of this nature was obtained by Breisford [19,20]. He experimentally verified the fact that pure predictions of cyclostationary meteorological data based on autoregressive techniques which employ periodically time-varying coefficients can be superior to similar predictions obtained with constant coefficients.

filters.²⁶

There are two dominant motives for this approach:

(1) In the past, time-varying systems were considerably more difficult--if not totally impractical--to implement compared with time-invariant systems, so that even if an optimum periodically time-varying filter was identified, it probably couldn't (or wouldn't) be built.

(2) Solutions for optimum (noncausal) filters for stationary processes are very simple to obtain, in direct contrast to the complete lack of any general solution methods for optimum (noncausal) filters for non-stationary processes. (The appropriateness of the noncausal optimum filter for communication problems will be discussed in Section 3.)

Thus, the tack taken in the past is very understandable. However, with the advent of integrated circuit technology, time-varying systems--particularly periodically time-varying systems--are becoming increasingly easier to implement. In fact, some such implementations are even more attractive in terms of design-and-construction effort and cost than classical time-invariant implementations based on resistors, capacitors, and inductors. With this new technology, the question of improvements in performance available through recognition of, and design for, the cyclic character of cyclostationary signals becomes very interesting and practical--and challenging, in light of the general lack of appropriate solution methods.

²⁶ The most predominant exception to this statement is the coherent demodulator-filter, for AM signals, which has been derived in various ways by many investigators.

2. The Orthogonality Condition

The necessary and sufficient condition which implicitly specifies the optimum (linear) filter was derived for the case of stationary ergodic processes and an infinite-memory filter by Wiener [56] in the 1940's. In 1950, the condition was derived for the case of stationary ergodic processes and a finite memory filter by Zadeh and Ragazzini [57], and in 1952 for the general case of nonstationary processes by Booton [58]. These classical results have been rederived, modified, and generalized by a number of authors using various tools of analysis such as the calculus of variations, directional derivative and gradient techniques in functional analysis, the Hilbert-space projection theorem, and others. We will present here a short derivation of a very general form of the necessary and sufficient condition for optimality by using the projection theorem stated below [29,59-61]:

Projection Theorem: Let H be a Hilbert space, and S a closed subspace of H . For every element $h \in H$ there exist a unique element $\hat{h} \in S$ such that

$$\|h - \hat{h}\|_H < \|h - s\|_H \quad \forall s \in S, s \neq \hat{h} \quad (4-1)$$

and

$$(h - \hat{h}, s)_H = 0 \quad \forall s \in S. \quad (4-2)$$

In this theorem \hat{h} is called the orthogonal projection of h onto S , and Eq. (4-2) is called the orthogonality condition. In words, this condition states that the difference between h and its orthogonal projection \hat{h} is orthogonal to all elements in the subspace S onto which h was projected.

A simple example of this theorem is illustrated in Fig. (4-2) where we have chosen the two-dimensional Euclidean space as our Hilbert space. In this space, the theorem is obvious from geometrical considerations.

Now, this theorem can be applied to many optimization problems simply by identifying the quantity to be minimized as the norm of the difference between a given element h and a variable (adjustable) element s in an appropriate Hilbert space H , and by identifying the constraints on the variable element as equivalent to the restriction of that element to an appropriate subspace S of the Hilbert space H . For the least-mean-square (LMS) estimation problem, we set up a separate (independent) optimization problem for each value of the time-index: For any specific value of time, say t , we identify H as H_{RV} the space of finite mean-square random variables (defined in Section 1 of Chapter II), and h as the transmitted signal process x evaluated at time t , and s as the output of the estimator which is to be optimized, so that minimizing the mean-squared estimation-error is the same as minimizing the norm

$$\|h-s\|_H^2 = \|x(t)-s(t)\|_{H_{RV}}^2 = E\{|x(t)-s(t)|^2\}, \quad t \in T_0$$

subject to

$$s(t) = L^{(t)} \cdot [y]$$

where y is the observed (received) signal process, and where $L^{(t)}$ is one element of a set of linear functionals, the collection of which

comprises the estimator transformation $L = \{L^{(t)}; t \in T_0\}$. Now, in order to apply the projection theorem, we need only identify the subspace $S_t \subset H_{RV}$ such that the restriction of L to a "permissible" set L is equivalent to the restriction of $s(t)$ to the subspace S_t for all values of $t \in T_0$. We first give a precise definition of the term "permissible":

DEFINITION: $L = \{L_i\}$ is defined to be a permissible set of linear transformations w.r.t. the observed process y , and $P = \{h_i(\cdot, \cdot)\}$ the permissible set of kernels (impulse-response functions), if and only if:

1. The sets P_t of kernel-sections $\{h_i(t, \cdot)\}$, corresponding to the linear functionals $L^{(t)}$, form Banach spaces with the $L^1(-\infty, \infty)$ - norm; i.e., the sets P_t form linear vector spaces which are complete w.r.t. the norm:

$$\|h_i(t, \cdot)\|_{L^1} = \int_{-\infty}^{\infty} |h_i(t, \tau)| d\tau < \infty.$$

2. If y contains a non-zero white component, then the kernel-sections form Banach spaces with the $L^2(-\infty, \infty)$ -norm:

$$\|h_i(t, \cdot)\|_{L^2} = \left(\int_{-\infty}^{\infty} h_i^2(t, \tau) d\tau \right)^{1/2} < \infty.$$

Note that P is also a linear vector space, and its elements are impulse-response functions of "bounded-input-bounded-output" stable systems [52] (because of the finite L^1 - norm). Note also that

although the finite L^2 -norm constraint may seem unnecessarily restrictive--in the sense that it excludes systems with direct connections from input to output--it is a practical constraint since it is necessary in order that the estimate of a finite mean-square process, based on observations containing a white component, be a finite mean-square process.

Now, we have the following theorem:

THEOREM(4-1): Let L be the permissible set of linear transformations corresponding to an arbitrary permissible set of kernels P , and let y be a random process which can be finite-mean-square, or white, or the sum of both (assumed uncorrelated). Then, the restriction of the transformation L to the set L is equivalent to the restriction of the random variables $\{s(t); t \in T_0\}$ to respective closed subspaces S_t of the Hilbert space H_{RV} , where

$$s(t) = L^{(t)} \cdot [y], \quad \{L^{(t)}; t \in T_0\} = L \in L,$$

Proof: By definition of L , P , P_t , we have the fact that for any transformation $L \in L$, there exist kernel-sections $h_i(t, \cdot) \in P_t$ such that

$$s(t) = L^{(t)} \cdot [y] = \int_{-\infty}^{\infty} h_i(t, \tau) y(\tau) d\tau \quad \forall t \in T_0.$$

Now,

$$\begin{aligned} E\{s^2(t)\} &= \iint_{-\infty}^{\infty} h_i(t, \tau) h_i(t, \sigma) k_{yy}(\tau, \sigma) d\tau d\sigma \\ &= \iint_{-\infty}^{\infty} h_i(t, \tau) h_i(t, \sigma) \hat{k}_{yy}(\tau, \sigma) d\tau d\sigma + \lambda \int_{-\infty}^{\infty} h_i^2(t, \tau) d\tau \end{aligned}$$

where \hat{k}_{yy} is the autocorrelation function of the finite mean-square component of y , and λ is the PSD of the white component. Thus, we have

$$E\{s^2(t)\} \leq \max_{\tau} \hat{k}_{yy}(\tau, \tau) \left[\int_{-\infty}^{\infty} |h_i(t, \sigma)| d\sigma \right]^2 + \lambda \int_{-\infty}^{\infty} h_i^2(t, \sigma) d\sigma < \infty$$

$$\forall t \in T_0$$

since $h_i(t, \cdot) \in L^1(-\infty, \infty)$ and since $h_i(t, \cdot) \in L^2(-\infty, \infty)$ if $\lambda \neq 0$.

Hence, $s(t) \in H_{RV} \forall t \in T_0$. Also, since P_t is a Banach space with both L^1 and L^2 norms, it is easily shown that the set of all $s(t)$ generated from y using all the elements of P_t is a closed linear vector space. This space of finite mean-square random variables is, then, a closed subspace of H_{RV} .

QED

Using the result of this last theorem, and the projection theorem, we obtain the following theorem which provides a necessary and sufficient condition for optimum linear filtering:

THEOREM(4-2): Let P be any permissible set of kernels. The filter with impulse-response function $h_0(\cdot, \cdot)$ minimizes the mean-squared error

$$e(t) = E\{(\hat{x}(t) - x(t))^2\} \quad \forall t \in T_0$$

where

$$\hat{x}(t) = \int_{-\infty}^{\infty} h_0(t, \tau) y(\tau) d\tau, \quad (4-3)$$

subject to $h_0(\cdot, \cdot) \in P$, if and only if

$$\iint_{-\infty}^{\infty} h_i(t, \tau) h_0(t, \sigma) k_{yy}(\tau, \sigma) d\tau d\sigma = \int_{-\infty}^{\infty} h_i(t, \tau) k_{yx}(\tau, t) d\tau \quad (4-4)$$

for every $h_i(\cdot, \cdot) \in P$, and for every $t \in T_0$, and the minimum mean-squared error is

$$e_0(t) = k_{xx}(t, t) - \int_{-\infty}^{\infty} h_0(t, \tau) k_{yx}(\tau, t) d\tau. \quad (4-5)$$

Proof: From the projection theorem (Eq. (4-2)), $\hat{x}(t)$ is the LMS estimate of $x(t) \in H_{RV}$, $\forall t \in T_0$, if and only if

$$E\{(\hat{x}(t) - x(t))s(t)\} = 0 \quad \forall s(t) \in S_t \subset H_{RV}, \quad \forall t \in T_0.$$

But, from Theorem (4-1),

$$S_t = \left\{ \int_{-\infty}^{\infty} h_i(t, \tau) y(\tau) d\tau; \quad h_i(t, \cdot) \in P_t \right\} \quad \forall t \in T_0. \quad (4-6)$$

Using this representation for S_t , and the formula (Eq. (4-3)) for $\hat{x}(t)$ in the above orthogonality condition yields the necessary and sufficient condition of Eq. (4-4). Furthermore, the minimum mean-squared error can be written as

$$e_0(t) = E\{(x(t) - \hat{x}(t))x(t)\} - E\{(x(t) - \hat{x}(t))\hat{x}(t)\},$$

and since $\hat{x}(t) \in S_t$, then the last term is zero. Now, using Eq. (4-3) we obtain the formula of Eq. (4-5).

QED

This necessary and sufficient condition of Eq. (4-4) is a considerable generalization of the integral equation frequently, but often inappropriately, referred to as the Wiener-Hopf equation for optimum filters. Several specific examples are given below:

(1) "Infinite-dimensional" filters: If the linear vector space P of permissible kernels is infinite-dimensional, then we will call the optimum filter in P an infinite-dimensional filter.²⁷ Typical constraints which lead to infinite-dimensional P 's are constraints on memory length. For example, we may define P such that

$$h_i(t, \tau) = 0, \forall t, \tau \ni t - \tau < T_{rm}, \text{ or } t - \tau > T_m, \text{ or } t, \tau \notin [T_i, T_f] = T_0.$$

for every $h_i(\cdot, \cdot) \in P$, where T_m is the memory length of the filter (amount of past input used to obtain the current output), T_{rm} is the reverse memory length of the filter (amount of future input used to obtain the current output--also, the amount of delay necessary to make the filter causal), and T_i, T_f are the initial and final times which define the interval over which the filter is to operate. In this case, our general condition for optimality becomes:

$$\int_{a(t)}^{b(t)} h_0(t, \sigma) k_{yy}(\sigma, \tau) d\sigma = k_{yx}(\tau, t), \begin{cases} \forall \tau \in [a(t), b(t)] \\ \forall t \in [T_i, T_f] \end{cases} \quad (4-7)$$

²⁷ Note that finite order linear dynamical systems are sometimes referred to as finite-dimensional systems [52], but are, under the above definition, infinite-dimensional. Furthermore note that the finite dimensional filters defined here do not necessarily have finite dimensional range spaces.

where

$$\begin{aligned} a(t) &\triangleq \max\{T_i, t-T_m\} \\ b(t) &\triangleq \min\{T_f, t+T_{rm}\}. \end{aligned} \quad (4-7)$$

As a specific example, if we choose $-T_i = T_f = T_m = \infty$, $T_{rm} = 0$, then P is the set of all causal filters, and if y and x are jointly WSS, then the above equation becomes a Wiener-Hopf equation.

(2) "Finite-dimensional" filters: If P is spanned by a finite set of kernels, then we call the optimum filter in P a finite-dimensional filter. For example, if P is M -dimensional, then every impulse-response function in P can be expressed as a linear time-varying combination of M basis kernels:

$$h(t, \tau) = \sum_{i=1}^M \alpha_i(t) h_i(t, \tau),$$

where the a-priori choice of the basis kernels defines the vector space P . Notice that the scalars in P are not real numbers as in P_t , but rather real time-functions. This is a result of the way in which P was constructed from $\{P_t; t \in T_0\}$.

For finite-dimensional filters, our optimality condition becomes:

$$\begin{aligned} \sum_{i=1}^M \alpha_{i0}(t) \iint_{-\infty}^{\infty} h_i(t, \tau) h_j(t, \sigma) k_{yy}(\tau, \sigma) d\tau d\sigma \\ = \int_{-\infty}^{\infty} h_j(t, \sigma) k_{yx}(\sigma, t) d\sigma \end{aligned} \quad (4-8)$$

for every $t \in T_0$, and for $j = 1, 2, \dots, M$, where $\{\alpha_{i0}(t)\}$ is the optimum set of weighting functions. Hence, the optimum M -dimensional filter has impulse-response function

$$h_o(t, \tau) = \sum_{i=1}^M \alpha_{io}(t) h_i(t, \tau),$$

where the M-vector of weights $\underline{\alpha}_o(t)$ is, from Eq. (4-8)

$$\underline{\alpha}_o(t) = [\underline{A}(t)]^{-1} \underline{b}(t) \quad (4-9)$$

where $\underline{A}(t)$ is the M x M matrix with elements:

$$a_{ij}(t) = \iint_{-\infty}^{\infty} h_i(t, \tau) h_j(t, \sigma) k_{yy}(\tau, \sigma) d\tau d\sigma \quad (4-10)$$

and $\underline{b}(t)$ is the M-vector with elements:

$$b_i(t) = \int_{-\infty}^{\infty} h_i(t, \tau) k_{yx}(\tau, t) d\tau. \quad (4-11)$$

Note that the basis filters can be chosen to have arbitrary memory length, and reverse memory length; they can be chosen to be arbitrarily time-varying filters, or causal time-invariant filters. They need only be stable filters as defined in the definition of P. Specific choices can be based on physical considerations of implementation, and on considerations of efficiency (minimization of the dimension necessary to reduce the mean-squared estimation-error below a prescribed level.) A formal realization of an arbitrary M-dimensional filter is shown in Fig. (4-3) as a parallel connection of M basis filters each followed by a multiplier.

In this chapter we will be primarily concerned with infinite dimensional noncausal filters (with infinite forward and reverse memory length) on the doubly infinite interval $(-\infty, \infty)$; i.e., we will make the choice of parameters: $T_m = T_{rm} = -T_1 = T_1 = \infty$, so that the general

necessary and sufficient condition for optimality becomes:

$$\int_{-\infty}^{\infty} h_0(t,s)k_{yy}(s,\tau)ds = k_{yx}(\tau,t), \quad \forall t,\tau \in (-\infty,\infty), \quad (4-12)$$

and the minimum mean-squared estimation error is given by

$$e_0(t) = k_{xx}(t,t) - \int_{-\infty}^{\infty} h_0(t,s)k_{yx}(s,t)ds, \quad \forall t \in (-\infty,\infty). \quad (4-13)$$

Before proceeding, we comment on our choice of the "integral-equation approach" to our linear least-mean-square estimation problems, as opposed to the "differential-equation approach", which incidentally can also be developed from the Projection Theorem. The type of LMS estimation problems considered here, like many other communication-systems-analysis problems, are most easily formulated in terms of integral transformations and impulse-response functions, whereas typical control-systems-analysis problems (and some communications problems) are most easily formulated in terms of differential transformations and state-evolution equations. For example, the various communication-signal formats and models for cyclostationary communications signals which were presented in Chapter II were developed in terms of subjecting elementary information bearing random processes (including multidimensional continuously indexed processes as well as quantized discretely indexed sequences) to systems (including transducers, modulators, multiplexors, and various signal processors) which were conveniently characterized as integral transformations with easily specified impulse-response functions. Equivalent state-equation descriptions of these signals and systems would, in most cases, be very cumbersome or impossible due to noncausality

and/or infinite-dimensionality. For example, the concept of an ideal bandlimiting filter is indispensable in many analyses of communications problems, (including data-sampling and frequency division multiplexing), yet is not amenable to state-variable description since the filter is noncausal; and the concept of an ideal delay-line is of great utility in many communications problems (including filtering of pulsed signals, and time-division multiplexing), but is an infinite dimensional (order) differential system and is therefore not amenable to state-variable description; and time-(and other index) scale transformations (which are useful for characterizing scanning operations, Doppler effects, signal compression, etc.) are easily described with impulse-response functions, but are formidable with state equations due to infinite-dimensionality; and so on.

Now, since the state-variable approach to LMS estimation problems--Kalman filtering--requires that all signals or processes be generated from white noise via a causal finite-order differential system [60,61], then it is of severely limited value for the problems of concern here. Furthermore, since we are specifically interested in noncausal filtering (with infinite reverse memory), then the Kalman filter (which must be causal with either zero or finite reverse memory) is of little value to us. In addition, "finite-dimensional" optimum filters (not necessarily finite order differential systems) are easily solved for with the integral-equation approach as discussed in this section, yet are excluded from the differential equation approach. Finally, we mention that our approach leads to explicit solutions for optimum filters and performances,

and useful insight into optimum time-varying signal processors, whereas the Kalman filtering approach leads to the Riccati matrix differential equation which almost always must be solved numerically.

One last comment: All results derived in this chapter using minimum-mean-squared-error as an optimality criterion are also valid when the optimality criterion is minimum-error-variance provided that all processes involved are replaced with their centered versions; e.g., x replaced with $x - m_x$.

3. Optimum Time-Invariant Filters for Stationary Processes

The optimum noncausal filter for wide-sense stationary processes is very simple to solve for, and yet is of significant theoretical and practical value. Three of the most outstanding reasons for its utility are:

- (1) The explicit solution gives useful insight into the nature of optimum filters.
- (2) A delayed version of the optimum noncausal filter can often be closely approximated by a causal filter.
- (3) The mean-squared estimation-error which results from the optimum noncausal filter is a useful performance bound since no causal filter can possibly perform any better.

The second fact above has useful application in communication problems since communication systems are often tolerant of, or insensitive to, the incorporation of a delay in the receiver. (Note, however, that this is far less frequently true for control problems, since control systems are often feedback systems, and are, by nature, very sensitive to the incorporation of delays, and can in fact become unstable when relatively small delays are included in the feedback loop.)

With this as motivation, we proceed to derive the optimum noncausal filter. Assuming that the transmitted and received signals are WSS, the necessary and sufficient condition for optimality (Eq. (4-12)) can be written as

$$\int_{-\infty}^{\infty} k_{yy}(\tau-s)h_0(t,s)ds = k_{xy}(\tau-t) \quad \forall t, \tau \in (-\infty, \infty).$$

But the solution to this equation depends only on the difference of its two arguments. Using this fact, and making a change of variables results in the equivalent condition for optimality:

$$\int_{-\infty}^{\infty} k_{yy}(t-\tau)h_0(\tau)d\tau = k_{xy}(t) \quad \forall t \in (-\infty, \infty),$$

which is a convolution. Hence, the Fourier transform of h_0 --the transfer function for the optimum filter--is given explicitly by the formula

$$H_0(f) = K_{xy}(f)/K_{yy}(f). \quad (4-14)$$

Using this formula in Eq. (4-13) and employing Parseval's relation [2] yields the following formula for the minimum mean-squared estimation-error:

$$e_0(t) = \int_{-\infty}^{\infty} \frac{K_{xx}(f)K_{yy}(f) - |K_{xy}(f)|^2}{K_{yy}(f)} df \quad \forall t \in (-\infty, \infty) \quad (4-15)$$

where K_{yx} is the cross PSD for y and x , and K_{yy} , K_{xx} are the PSD's for y , x .

If we consider the special class of problems where the channel is composed of a time-invariant dispersion with transfer function G and additive noise n , assumed uncorrelated with the signal x , then the optimum filter has transfer function given by

$$H_0(f) = \frac{G^*(f)K_{xx}(f)}{|G(f)|^2 K_{xx}(f) + K_{nn}(f)} \quad (4-16)$$

where K_{nn} is the PSD for the additive noise. This filter, frequently referred to as the Wiener filter, results in a minimum mean-squared

estimation-error which is given by the formula

$$e_o = \int_{-\infty}^{\infty} K_{xx}(f) |1-G(f)H_o(f)|^2 df + \int_{-\infty}^{\infty} K_{nn}(f) |H_o(f)|^2 df. \quad (4-17)$$

In this expression, the first term represents error due to imperfect compensation for channel dispersion, while the second term represents error due to the filtered noise remaining at the output. A more compact formula for the error which results directly from Eq. (4-15), is

$$e_o = \int_{-\infty}^{\infty} \frac{K_{nn}(f)K_{xx}(f)}{|G(f)|^2 K_{xx}(f) + K_{nn}(f)} df \quad (4-18)$$

and reduces to

$$e_o = \int_{-\infty}^{\infty} \frac{N_o K_{xx}(f)}{N_o + K_{xx}(f)} df = N_o h_o(0) \quad (4-19)$$

when there is no dispersion ($G(f) = 1$) and the additive noise is white with PSD equal to N_o . Although Eq. (4-19) is simple in form, it can present difficulties in terms of numerical computation because of the infinite interval of integration. With this in mind, we present the following error bounds which require evaluation of an integral only on a finite interval:

THEOREM(4-3): The minimum mean-squared estimation-error which results from the Wiener filter for the dispersionless channel is contained within the ± 1.5 dB bounds:

$$B/2 \leq e_o \leq B, \quad B = k_{xx}(0) - 2 \int_0^{f_o} [K_{xx}(f) - K_{nn}(f)] df \quad (4-20)$$

provided that the signal PSD dominates the noise PSD for all frequencies

less than f_0 , and vice-versa for all frequencies greater than f_0 :

$$K_{xx}(f) \begin{cases} \geq K_{nn}(f), & |f| \leq f_0 \\ \leq K_{nn}(f), & |f| \geq f_0 \end{cases}$$

Proof: By hypothesis,

$$\begin{aligned} \int_0^{f_0} K_{nn}(f) df &= \int_{-f_0}^{f_0} \frac{K_{nn}(f)K_{xx}(f)}{K_{xx}(f) + K_{nn}(f)} df \leq \int_{-f_0}^{f_0} \frac{K_{nn}(f)K_{xx}(f)}{K_{nn}(f) + K_{xx}(f)} df \\ &\leq \int_{-f_0}^{f_0} K_{nn}(f) df = 2 \int_0^{f_0} K_{nn}(f) df, \end{aligned}$$

and

$$\begin{aligned} \int_{f_0}^{\infty} K_{xx}(f) df &= 2 \int_{f_0}^{\infty} \frac{K_{nn}(f)K_{xx}(f)}{K_{nn}(f) + K_{xx}(f)} df \\ &\leq 2 \int_{f_0}^{\infty} \frac{K_{nn}(f)K_{xx}(f)}{K_{nn}(f) + K_{xx}(f)} df \leq 2 \int_{f_0}^{\infty} K_{xx}(f) df \end{aligned}$$

Adding these two strings of inequalities, and using the fact that

$$2 \int_{f_0}^{\infty} K_{xx}(f) df = k_{xx}(0) - 2 \int_0^{f_0} K_{xx}(f) df,$$

yields the bounds of Eq. (4-20).

QED

4. Optimum Time-Invariant Filters for Cyclostationary Processes

As discussed in the introductory section of this chapter, an optimum (in a sense to be defined below) time-invariant filter for cyclostationary processes can be obtained simply by replacing the given transmitted and received T-CS processes x, y with their stationary versions \tilde{x}, \tilde{y} (Theorem (2-15))--which have realizations of the form $\tilde{x}(t) = x(t+\theta), \tilde{y}(t) = y(t+\theta)$, where θ is a random phase-variable which is uniformly distributed over the period $[-T/2, T/2]$ --and then solving for the optimum filter for these stationary processes. The auto- and cross-correlation functions for these stationary processes \tilde{x}, \tilde{y} are (Theorem (2-16)):

$$\begin{aligned}\tilde{k}_{yy}(\tau) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{yy}(t+\tau, t) dt \\ \tilde{k}_{xy}(\tau) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{xy}(t+\tau, t) dt\end{aligned}\quad (4-21)$$

so that the optimum time-invariant filter has transfer function given by (Eq. (4-14))

$$\tilde{H}_0(f) = \tilde{k}_{xy}(f) / \tilde{k}_{yy}(f), \quad (4-22)$$

where $\tilde{k}_{xy}, \tilde{k}_{yy}$ are the Fourier transforms of $\tilde{k}_{xy}, \tilde{k}_{yy}$ and are therefore PSD's for the stationary processes \tilde{x}, \tilde{y} . However, the mean-squared estimation-error resulting from this filter will not be given by Eq. (4-15), since \tilde{e}_0 is, by definition,

$$\tilde{e}_0(t) = E\{(x(t) - \hat{x}(t))^2\} \neq E\{(\tilde{x}(t) - \hat{x}(t))^2\}$$

and is, in fact, a T-periodic function rather than a constant. On the other hand, the time-averaged value of $\tilde{e}_0(t)$ is indeed given by the formula in Eq. (4-15):

$$\langle \tilde{e}_0 \rangle \triangleq \frac{1}{T} \int_{-T/2}^{T/2} \tilde{e}_0(t) dt = \int_{-\infty}^{\infty} \frac{\tilde{K}_{xx}(f)\tilde{K}_{yy}(f) - |\tilde{K}_{xy}(f)|^2}{\tilde{K}_{yy}(f)} df. \quad (4-23)$$

Furthermore, the suboptimum filter with transfer function \tilde{H}_0 (sub- because it ignores the cyclic character of the cyclostationary processes) is actually optimum subject to the constraint of time-invariance; i.e., it minimizes the time-averaged value of the periodically time-varying mean-squared estimation-error. This result is stated here as a theorem:

THEOREM(4-4): The optimum filter, given by Eq. (4-22), for the stationarized versions \tilde{x} , \tilde{y} of the cyclostationary processes x , y is time-invariant and is the same as the filter which minimizes the time-averaged value of the periodic mean-squared estimation-error $\langle \tilde{e} \rangle$, and the minimum error $\langle \tilde{e}_0 \rangle$ is given by Eq. (4-23).

Proof: By definition,

$$\begin{aligned} \langle \tilde{e} \rangle &= \frac{1}{T} \int_{-T/2}^{T/2} E\left\{ \left[x(t) - \int_{-\infty}^{\infty} h(\tau) y(t-\tau) d\tau \right]^2 \right\} dt \\ &= \frac{1}{T} \int_{-T/2}^{T/2} \left[k_{xx}(t, t) - 2 \int_{-\infty}^{\infty} h(\tau) k_{xy}(t, t-\tau) d\tau \right. \\ &\quad \left. + \iint_{-\infty}^{\infty} h(\sigma) h(\tau) k_{yy}(t-\sigma, t-\tau) d\sigma d\tau \right] dt \end{aligned}$$

$$\begin{aligned}
&= \hat{k}_{xx}(0) - 2 \int_{-\infty}^{\infty} h(\tau) \hat{k}_{xy}(t-(t-\tau)) d\tau + \iint_{-\infty}^{\infty} h(\sigma) h(\tau) \hat{k}_{yy}((t-\sigma)-(t-\tau)) d\sigma d\tau \\
&= E\{[\hat{x}(t) - \int_{-\infty}^{\infty} h(\tau) \hat{y}(t-\tau) d\tau]^2\}.
\end{aligned}$$

But this last expression is the mean-squared estimation-error for the problem with transmitted and received signals \hat{x} , \hat{y} and filter-impulse-response function h , and is therefore minimized when $h = \hat{h}_0$, the inverse Fourier transform of \hat{H}_0 given by Eq. (4-22). Furthermore, the minimum value of this time-averaged error is

$$\langle \tilde{e}_0 \rangle = E\{(\hat{x}(t) - \hat{x}(t))^2\} = E\{[\hat{x}(t) - \int_{-\infty}^{\infty} \hat{h}_0(\tau) \hat{y}(t-\tau) d\tau]^2\}$$

and is therefore given by Eq. (4-23).

QED

If we consider the special class of problems where the channel consists of a time-invariant dispersion with transfer function G , and additive WSS noise with PSD K_{nn} , then paralleling the results in the preceding section, we have the following formula's for the transfer function for the optimum time-invariant filter, and for the resulting minimum time-averaged error.

$$\hat{H}_0(f) = \frac{G^*(f) \hat{k}_{xx}(f)}{|G(f)|^2 \hat{k}_{xx}(f) + K_{nn}(f)} \quad (4-24)$$

$$\langle \tilde{e}_0 \rangle = \int_{-\infty}^{\infty} \frac{K_{nn}(f) \hat{k}_{xx}(f)}{|G(f)|^2 \hat{k}_{xx}(f) + K_{nn}(f)} df \quad (4-25)$$

and for white noise and no dispersion

$$\hat{H}_0(f) = \frac{\hat{k}_{xx}(f)}{\hat{k}_{xx}(f) + N_0} \quad (4-26)$$

$$\langle \tilde{e}_0 \rangle = \int_{-\infty}^{\infty} \frac{N_0 \tilde{K}_{xx}(f)}{\tilde{K}_{xx}(f) + N_0} df = N_0 \tilde{h}_0(0). \quad (4-27)$$

These last four formulas fully characterize the optimum time-invariant filters and corresponding time-averaged estimation-errors for the specific illustrative problems that we will consider in the last section (6) of this chapter.

5. Optimum Time-Varying Filters for Cyclostationary Processes

a) Introduction. In this section we present a number of general methods of solution for optimum time-varying noncausal linear filters for cyclostationary processes. In contrast to the optimum time-invariant filters presented in the previous section, these filters are not constrained to be time-invariant and, in fact, minimize the mean-squared estimation-error for every value of time thereby taking full account of the cyclic character of the cyclostationary processes. As claimed in the introductory section, these optimum periodically time-varying linear filters result in lower mean-squared estimation-errors, as will be demonstrated via a number of specific solutions in the following section.

In order to characterize the relative performance of optimum time-invariant and optimum time-varying filters with a single number, we choose as a performance index I the ratio of the minimum time-averaged mean-squared estimation-error $\langle \tilde{e}_0 \rangle$, resulting from the optimum time-invariant filter, to the time-averaged value of the minimum mean-squared estimation-error $\langle e_0 \rangle$ resulting from the optimum time-varying filter:

$$I \frac{\Delta \langle \tilde{e}_0^2 \rangle}{\langle e_0 \rangle} \geq 1, \quad (4-28)$$

where $\langle \tilde{e}_0^2 \rangle$ is given by Eqs. (4-23, 25, 27), and where $\langle e_0 \rangle$ is from Eq. (4-13),

$$\langle e_0 \rangle = \frac{1}{T} \int_{-T/2}^{T/2} [k_{xx}(t, t) - \int_{-\infty}^{\infty} h_0(t, s) k_{yx}(s, t) ds] dt \quad (4-29)$$

where h_0 is the impulse-response function of the optimum time-varying filter and is the solution of the integral equation (Eq. (4-12)):

$$\int_{-\infty}^{\infty} h_0(t, s) k_{yy}(s, \tau) ds = k_{yx}(\tau, t) \quad \forall t, \tau \in (-\infty, \infty). \quad (4-30)$$

If we consider the special class of problems where the channel is a time-invariant dispersion with impulse-response function g and additive independent WSS noise with autocorrelation function k_{nn} , then the above condition for optimality becomes

$$\int_{-\infty}^{\infty} h_0(t, s) [k_{zz}(s, \tau) + k_{nn}(s, \tau)] ds = k_{zx}(\tau, t) \quad \forall t, \tau \in (-\infty, \infty)$$

where,

$$\begin{aligned} k_{zz}(t, \tau) &= \iint_{-\infty}^{\infty} g(t-\sigma) g(\tau-s) k_{xx}(\sigma, s) d\sigma ds \\ k_{zx}(t, \tau) &= \int_{-\infty}^{\infty} g(t, \sigma) k_{xx}(\sigma, \tau) d\sigma. \end{aligned} \quad (4-31)$$

For the case of white noise and no dispersion, the formula for the time-averaged minimum error reduces to

$$\langle e_0 \rangle = \frac{N_0}{T} \int_{-T/2}^{T/2} h_0(t, t) dt \quad (4-32)$$

where the optimum impulse-response function is now the solution of the integral equation

$$\int_{-\infty}^{\infty} h_o(t,s)k_{xx}(s,\tau)ds + N_o h_o(t,\tau) = k_{xx}(t,\tau) \quad \forall t,\tau \in (-\infty,\infty) \quad (4-33)$$

For this class of additive-white-noise problems, our performance index takes the form

$$I = \frac{\tilde{h}_o(0)}{\frac{1}{T} \int_{-T/2}^{T/2} h_o(t,t)dt} \quad (4-34)$$

In the remaining subsections, we present four different methods for solution of the integral equation Eq. (4-30) (and its special forms Eqs. (4-31,33)) which implicitly specifies the optimum filter and minimum estimation-error. Although these solutions are not as simple to obtain as were the solutions for optimum time-invariant filters, they are still of significant practical and theoretical value for the same reasons cited in Section 3:

- (1) The explicit solution gives useful insight into the nature of optimum time-varying filters.
- (2) A delayed version of the optimum noncausal filter can often be closely approximated by a causal filter.
- (3) The mean-squared estimation-error which results from the optimum noncausal filter is a useful performance bound, since no causal filter can possibly perform better.

b) Solutions based on translation series representations. We begin by considering the general LMS estimation problem where the channel is arbitrary but where the transmitted and received signals are jointly T-CS. If we assume, for generality, that y is composed of a colored

component w and an uncorrelated white component with PSD λ , then the optimality condition of Eq. (4-30) becomes

$$\int_{-\infty}^{\infty} h_0(t,s) k_{ww}(s,\tau) ds + \lambda h_0(t,\tau) = k_{xw}(t,\tau) \quad \forall t, \tau \in (-\infty, \infty), \quad (4-35)$$

But this integral equation is of the same form as Eq. (3-39) of Section 2f in Chapter III (except for the juxtaposition of the kernels h_0, k_{ww}).

Hence, Theorem (3-2) gives us the following solution:

THEOREM(4-5): The impulse-response function for the optimum filter for the general LMS estimation problem where the transmitted and received signals x, y are jointly T-CS and where y is composed of a colored component w and an uncorrelated white component with PSD λ has the representation:

$$h_0(t,\tau) = \sum_{n,m=-\infty}^{\infty} \sum_{p,q=1}^{\infty} H_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT) \quad (4-36)$$

where $\{\phi_p\}$ is any CON set on $L^2[0,T]$ and where the sequence of matrices $\{\underline{H}_r\}$ is given by the inverse z-transform

$$\underline{H}_r = T \int_{-1/2T}^{1/2T} \underline{H}(f) e^{j2\pi r T f} df \quad (4-37)$$

where the matrix of functions $\underline{H}(f)$ is given by

$$\underline{H}(f) = \underline{G}(f) [\underline{D}(f) + \lambda \underline{I}]^{-1} \quad (4-38)$$

where the matrices $\underline{D}, \underline{G}$ are the z-transforms

$$\begin{aligned} \underline{D}(f) &= \sum_{r=-\infty}^{\infty} \underline{D}_r e^{-j2\pi r T f} \\ \underline{G}(f) &= \sum_r \underline{G}_r e^{-j2\pi r T f} \end{aligned} \quad (4-39)$$

and where the elements of the sequences of matrices \underline{D}_T , \underline{G}_T are given by

$$\begin{aligned} D_r^{pq} &= \int_0^T \int_0^T k_{ww}(t+rT, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau \\ G_r^{pq} &= \int_0^T \int_0^T k_{xw}(t+rT, \tau) \phi_p^*(t) \phi_q(\tau) dt d\tau, \end{aligned} \quad (4-40)$$

and are the correlation matrices corresponding to the translation series representations (TSR's) for x and w :

$$\begin{aligned} D_{n-m}^{pq} &= E\{d_{np} d_{mq}^*\} \\ G_{n-m}^{pq} &= E\{a_{np} d_{mq}^*\} \end{aligned} \quad (4-41)$$

where

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{\infty} \sum_{p=1}^{\infty} a_{np} \phi_p(t-nT) \\ w(t) &= \sum_n \sum_p d_{np} \phi_p(t-nT). \end{aligned} \quad (4-42)$$

Furthermore, the time-averaged value of the minimum mean-squared estimation-error is given by the formula:

$$\langle e_o \rangle = \int_{-1/2T}^{1/2T} \text{trace}[\underline{A}(f) - \underline{G}(f) [\underline{D}(f) + \lambda \underline{I}]^{-1} \underline{G}^*(f)] df \quad (4-43)$$

where $\underline{A}(f)$ is the correlation matrix corresponding to the TSR for x (defined as in Eq. (4-39,40) for w).

Proof: The validity of the solution Eq. (4-36) to the integral equation Eq. (4-35) follows directly from Theorem (3-2) of Chapter III. The formula Eq. (4-43) can be derived from Eq. (4-29) as follows: Substituting Eq. (4-36) for h_o , and the TSR's for k_{xx} and k_{xy} into Eq. (4-29) yields:

$$\begin{aligned}
\langle e_o \rangle &= \frac{1}{T} \int_0^T \left[\sum_{n,m} \sum_{p,q} A_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(t-mT) \right. \\
&\quad \left. - \int_{-\infty}^{\infty} \sum_{n,m} \sum_{p,q} H_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT) \cdot \right. \\
&\quad \left. \sum_{n',m'} \sum_{p',q'} G_{n'-m'}^{p'q'} \phi_{p'}^*(t-n'T) \phi_{q'}(s-m'T) ds \right] dt \\
&= \frac{1}{T} \sum_p [A_o^{pp} - \sum_m \sum_q H_m^{pq} G_m^{*pq}] \\
&= \int_{-1/2T}^{1/2T} \text{trace}[\underline{A}(f) - \underline{H}(f)\underline{G}^{*'}(f)] df.
\end{aligned}$$

Now, substituting Eq. (4-38) for $\underline{H}(f)$ yields Eq. (4-43).

QED

As discussed in Section 2g of Chapter III, a linear system whose impulse-response function has an M^{th} order TSR, as does h_o in this theorem (with $M = \infty$), can be realized with the structure shown in Figure (3-9) and reproduced here as Figure (4-4a). Although the number (M) of paths in this optimum filter for CS processes will in general be infinite, there are a number of interesting cases where M is finite, as discussed in Sec. 2f of Chapter III and as illustrated in the last section of this chapter. Furthermore, since any CS process in $H_{CS(T)}^{(-\infty, \infty)}$ can be arbitrarily closely approximated in norm by a finite order TSR (as discussed in Section 2a of Chapter III), then the performance of a truncated (to M paths) optimum filter can be made arbitrarily close to the optimum performance by choosing the finite number (M) of paths large enough.

An interesting interpretation of this optimum filter follows directly from the recognition that the bank of input filters and switches comprises the process-resolution operation (as discussed in Sec. 2c of Chapter III), and that the bank of output filters and the summing device comprise the process-reconstruction operation, and finally that the transfer function $\underline{H}(f)$ (Eq. (4-38)) for the matrix of sampled-data filters which is inserted between the resolution and reconstruction systems is a matrix version of the transfer function $H_0(f)$ (Eq. (4-14)) for the optimum time-invariant filter for stationary processes, in that $\underline{G}(f)$ in Eq. (4-38) is the cross-correlation matrix for the representors of x , y --paralleling $K_{xy}(f)$ of Eq. (4-14)--and $[\underline{D}(f) + \lambda \underline{I}]$ in Eq. (4-38) is the autocorrelation matrix for the representors of y --paralleling $K_{yy}(f) = K_{ww}(f) + \lambda$ of Eq. (4-14). Thus, we see that the optimum time-varying filter of Theorem (4-5) for CS processes resolves the received CS process y into a multiplicity of jointly WSS sequences, subjects these to an optimum time-invariant sampled-data filter, then reconstructs from these filtered sequences, the optimally filtered continuous-time estimate \hat{x} of the transmitted signal x . Hence, by employing TSR's we convert the problem of LMS estimation of a scalar CS process to the problem of LMS estimation of a vector of jointly WSS sequences. This could have been done in a non-rigorous suboptimum way simply by periodically sampling the received signal, and solving for the LMS estimates of the corresponding samples of the transmitted signal.

The optimum filter derived in Theorem (4-5) can also be realized with the alternate structure shown in Fig. (4-4b) where the filter-

sampler combinations at the input have been replaced with multiplier-integrators, and where the impulse-driven pulse-generating filters at the output have been replaced with pulse-amplitude modulators, and where the feedback-delay device (shown in Fig. (3-3)), which has an effect equivalent to periodically dumping the integrator, has been included in the sampled data filter by modifying its transfer function to include the factor $1/(1 + e^{-j2\pi fT})$.

As discussed in Chapter III, some CS signals which are used in communication systems admit finite order TSR's, but with basis functions which are not necessarily duration limited to $[0, T]$, nor even orthogonal. With this in mind, we present the following extension of Theorem (4-5):

THEOREM(4-6): If the transmitted signal x , and the colored component w of the received signal y admit M^{th} order TSR's

$$\begin{aligned} x(t) &= \sum_{p=1}^M \sum_{n=-\infty}^{\infty} a_{np} \phi_p(t-nT) \\ w(t) &= \sum_{p=1}^M d_{np} \theta_p(t-nT) \quad \forall t \in (-\infty, \infty) \end{aligned} \quad (4-44)$$

where the functions in the sets $\{\phi_p\}$, $\{\theta_p\}$ are arbitrary, then the optimum impulse-response function admits the representation

$$h_o(t, \tau) = \sum_{p,q=1}^M \sum_{n,m=-\infty}^{\infty} H_{n-m}^{pq} \phi_p(t-nT) \theta_q^*(\tau-mT) \quad (4-45)$$

where $\{H_{n-m}^{pq}\}$ are defined as in Eq. (4-37), where

$$\underline{H}(f) = \underline{G}(f) [\underline{E}(f) \underline{D}(f) + \lambda \underline{1}]^{-1} \quad (4-46)$$

where \underline{G} , \underline{D} are defined as in Eqs. (4-39), (4-41), and where \underline{E} is the z-transform of the matrix of sequences

$$E_{n-m}^{pq} = \int_{-\infty}^{\infty} \theta_p^*(t-nT) \theta_q(t-mT) dt. \quad (4-47)$$

Furthermore, the time-averaged value of the minimum mean-squared estimation-error is given by the formula:

$$\langle e_o \rangle = \int_{-1/2T}^{1/2T} \text{trace}[\underline{B}(f)\underline{A}(f) - \underline{B}(f)\underline{G}(f) [\underline{E}(f)\underline{D}(f) + \lambda \underline{I}]^{-1} \underline{E}(f)\underline{G}^*(f)] df \quad (4-48)$$

where \underline{B} is the z-transform of the matrix of sequences

$$B_{n-m}^{pq} = \int_{-\infty}^{\infty} \phi_p^*(t-nT) \phi_q(t-mT) dt. \quad (4-49)$$

The proof of this theorem follows directly that of Theorems (3-2), (4-5).

The optimum filter given in this theorem has the same structure as that given in the previous theorem and shown in Fig. (4-4), except that the input filters have transfer functions $\{\theta_p^*\}$ rather than $\{\phi_p^*\}$.

Another interesting solution for a common formulation of the LMS estimation problem is given in the following theorem:

THEOREM(4-7): Consider the class of LMS estimation problems where the channel consists of a time-invariant dispersion with transfer function G , and additive WSS noise with PSD K_{nn} , and where the transmitted signal x is a T-CS process which admits an M^{th} order TSR:

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M a_{np} \phi_p(t-nT)$$

For this class of problems, the impulse-response function for the optimum time-varying filter admits the M^{th} order TSR:

$$h_o(t, \tau) = \sum_{m, n=-\infty}^{\infty} \sum_{p, q=1}^M H_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(\tau-mT) \quad (4-50)$$

where θ_q is the inverse Fourier transform of the function

$$\theta_q(f) = G(f) \phi_q(f) / K_{nn}(f) \quad (4-51)$$

and where the sequence of $M \times M$ matrices $\{\underline{H}_r\}$ is given by the inverse z-transform

$$\underline{H}_r = T \int_{-1/2T}^{1/2T} \underline{H}(f) e^{j2\pi r T f} df$$

where the $M \times M$ matrix of functions $\underline{H}(f)$ is given by

$$\underline{H}(f) = \underline{A}(f) [\underline{C}(f) \underline{A}(f) + \underline{I}]^{-1} \quad (4-52)$$

where the matrices $\underline{A}(f)$, $\underline{C}(f)$ are the z-transforms

$$\underline{A}(f) = \sum_{r=-\infty}^{\infty} \underline{A}_r e^{-j2\pi r T f}$$

$$\underline{C}(f) = \sum_r \underline{C}_r e^{-j2\pi r T f}$$

and where the sequences of matrices \underline{A}_r , \underline{C}_r are given by the formulas

$$A_r^{pq} = E\{a_{(n+r)p} a_{nq}^*\}$$

$$C_r^{pq} = \int_{-\infty}^{\infty} \frac{|G(f)|^2 \phi_p^*(f) \phi_q(f)}{K_{nn}(f)} e^{j2\pi r T f} df; \quad (4-53)$$

Furthermore, the time-averaged value of the minimum estimation-error is given by the formula

$$\langle e_o \rangle = \int_{-1/2T}^{1/2T} \text{trace}[\underline{B}(f)\underline{A}(f)[\underline{C}(f)\underline{A}(f) + \underline{I}]^{-1}]df \quad (4-54)$$

where \underline{B} is defined in Eq. (4-49).

Proof: Substituting Eq. (4-50) for h_o and the TSR of Eq. (3-2) for k_{xx} into the condition for optimality-Eq. (4-31) yields:

$$\sum_{n,m} \sum_{p,q=1}^M \phi_p(t-nT) \left[\sum_{n',m'} \sum_{p',q'} A_{m'-m}^{q'q} H_{n-n'}^{pp'} \int_{-\infty}^{\infty} \theta_p^*(\sigma-n'T) \tilde{\phi}_q(\sigma-m'T) d\sigma \right. \\ \left. \tilde{\phi}_q^*(\tau-mT) + H_{n-m}^{pq} [\theta_q^* k_{nn}] (\tau-mT) - A_{n-m}^{pq} \tilde{\phi}_q^*(\tau-mT) \right] = 0 \quad \forall t, \tau \in (-\infty, \infty)$$

where

$$\tilde{\phi}_q(t) \triangleq [\phi_q \otimes g](t) = \int_{-\infty}^{\infty} \phi_q(t-\tau) g(\tau) d\tau.$$

But since $\theta_q^* k_{nn} = \tilde{\phi}_q^*$, then the above optimality condition can be written as

$$\sum_{n,m} \sum_{p,q=1}^M \left[\sum_{n',m'} \sum_{p',q'} H_{n-n'}^{pp'} C_{n'-m'}^{p'q'} A_{m'-m}^{q'q} + H_{n-m}^{pq} - A_{n-m}^{pq} \right] \cdot \\ \phi_p(t-nT) \tilde{\phi}_q^*(\tau-mT) = 0 \quad \forall t, \tau,$$

and will be satisfied if the composite coefficients of the time-functions are zero for every value of n, m, p, q . This condition on the coefficients can be written (using the new variables $i = m' - m$, $j = n - n'$, $k = n - m$) as:

$$\sum_{i,j=-\infty}^{\infty} \sum_{p',q'=1}^M H_j^{pp'} C_{k-i-j}^{p'q'} A_i^{q'q} + H_k^{pq} - A_k^{pq} = 0 \quad \forall k, q, p,$$

where

$$C_r^{pq} \triangleq \int_{-\infty}^{\infty} \theta_p^*(\sigma-rT) \tilde{\phi}_q(\sigma) d\sigma. \quad (4-55)$$

But this quadruple sum is simply the composition of a discrete convolution with another discrete convolution with a finite matrix product with another finite matrix product. Recognizing this, and taking z-transforms, yields the equivalent condition:

$$\underline{H}(f)\underline{C}(f)\underline{A}(f) + \underline{H}(f) - \underline{A}(f) = \underline{0} \quad \forall f \in (-\infty, \infty),$$

which is satisfied by the formula for \underline{H} given in the theorem. Furthermore using Parseval's relation in Eq. (4-55) yields Eq. (4-53) for the sequence of matrices C_r^{pq} . Now, the formula of Eq. (4-54) for the minimum estimation-error can be obtained from the general formula of Eq. (4-29) by substituting the TSR's for k_{xx} , h_o , k_{xy} into the latter to obtain:

$$\langle e_o \rangle = \sum_{p,q=1}^M \sum_{r=-\infty}^{\infty} B_{-r}^{qp} [A_r^{pq} - \sum_{n,m} \sum_{p',q'}^M H_{r-(n-m)}^{pp'} C_{n-m}^{p'q'} A_m^{q'q'}].$$

This number is the trace of the zeroth element in the sequence of matrices

$$\{\underline{B}\} \otimes \{\{\underline{A}\} - \{\underline{H}\} \otimes \{\underline{C}\} \otimes \{\underline{A}\}\}$$

where \otimes denotes the composition of discrete convolution and matrix product. Taking the z-transform yields

$$\underline{B}(f) [\underline{A}(f) - \underline{H}(f)\underline{C}(f)\underline{A}(f)] = \underline{B}(f)\underline{H}(f).$$

Hence, $\langle e_o \rangle$ is the trace of the zeroth element of the inverse z-transform of this matrix of functions, and is given by Eq. (4-54).

QED

The optimum filter given in this theorem has the same structure as that shown in Fig. (4-4), except that the input filters have transfer

functions given by

$$\theta_p^*(f) = G^*(f)\phi_p^*(f)/K_{nn}(f)$$

rather than ϕ_p^* , and are therefore matched filters for dispersed versions of the pulses $\{\phi_p\}$ in colored noise [2].

If we consider the class of optimum filtering problems where the channel consists of nothing more than additive white noise with PSD N_0 , then it is easily shown that the formulas of Theorems (4-5), (4-6), (4-7) for the optimum impulse-response function all reduce to the simplified versions:

$$h_o(t, \tau) = \sum_{p,q=1}^M \sum_{n,m=-\infty}^{\infty} H_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(\tau-mT) \quad (4-56)$$

where the z-transforms of the matrix of sequences $\{H_r^{pq}\}$ is given by the formula

$$\underline{H}(f) = \underline{A}(f) [\underline{E}(f) \underline{A}(f) + N_0 \underline{I}]^{-1} \quad (4-57)$$

where \underline{E} is the z-transform of the matrix of sequences

$$E_r^{pq} = \int_{-\infty}^{\infty} \phi_p^*(t-rT) \phi_q(t) dt = \int_{-\infty}^{\infty} \phi_p^*(f) \phi_q(f) e^{j2\pi r T f} df.$$

Also, the formulas for the time-averaged value of the minimum estimation-error given in the previous three theorems reduce to the formula

$$\langle e_o \rangle = N_0 \int_{-1/2T}^{1/2T} \text{trace}[\underline{E}(f) \underline{A}(f) [\underline{E}(f) \underline{A}(f) + N_0 \underline{I}]^{-1}] df. \quad (4-58)$$

Note the parallel to the formulas for the optimum time-invariant filter transfer function H_0 and for the resultant minimum value of the time-averaged estimation-error $\langle \tilde{e}_0 \rangle$:

$$H_0(f) = \frac{\frac{1}{T} \underline{\phi}'(f) \underline{A}(f) \underline{\phi}^*(f)}{N_0 + \frac{1}{T} \underline{\phi}'(f) \underline{A}(f) \underline{\phi}^*(f)} \quad (4-59)$$

$$\langle \tilde{e}_0 \rangle = N_0 \int_{-\infty}^{\infty} \frac{\frac{1}{T} \underline{\phi}'(f) \underline{A}(f) \underline{\phi}^*(f)}{N_0 + \frac{1}{T} \underline{\phi}'(f) \underline{A}(f) \underline{\phi}^*(f)} df = N_0 \tilde{h}_0(0). \quad (4-60a)$$

These latter two formulas result from substituting Eq. (3-7) for \tilde{K}_{xx} into the general formulas of Eqs. (4-26), (4-27).

Furthermore, if the $\{\phi_p(t-nT)\}$ are doubly orthogonal, in the sense that

$$\int_{-\infty}^{\infty} \phi_p(t-nT) \phi_q^*(t-mT) dt = \delta_{nm} \delta_{pq},$$

as in Theorem (4-5), then $\underline{E}(f) = \underline{I}$, and the formula of Eq. (4-57) reduces to

$$\underline{H}(f) = \underline{A}(f) [\underline{A}(f) + N_0 \underline{I}]^{-1} \quad (4-61a)$$

and $\underline{H}(f)$ is the transfer function for a matrix sampled-data Wiener filter for jointly WSS sequences in additive white noise. For this case the formula for the minimum error also reduces to

$$\langle \tilde{e}_0 \rangle = N_0 \int_{-1/2T}^{1/2T} \text{trace}[\underline{A}(f) [\underline{A}(f) + N_0 \underline{I}]^{-1}] df = \frac{N_0}{T} \text{trace}[\underline{H}_0]. \quad (4-62a)$$

As discussed in Subsection 2f of Chapter III, there are various CS signals of interest whose TSR correlation matrices are diagonal. For example, if the autocorrelation function satisfies the condition

$$k_{xx}(t+rT, \tau) = \alpha_r k_{xx}(t, \tau), \quad \forall t, \tau \in [0, T],$$

then from Theorem (3-3) of Chapter III the correlation matrix $\underline{A}(f)$ for the Karhunen-Loeve TSR is diagonal:

$$A_{pq}(f) = A_0^{pp} \delta_{pq} \sum_r \alpha_r e^{-j2\pi r T f}.$$

When $\underline{A}(f)$ is diagonal, the transfer function for the matrix of sample-data filters given by Eq. (4-61a) reduces to

$$H_{pq}(f) = \frac{\delta_{pq} A_{pp}(f)}{N_o + A_{pp}(f)} \quad (4-61b)$$

and is diagonal. Hence, the optimum filter structures shown in Figs. (4-4a), (4-4b) consist of a parallel connection of independent paths, as shown in Fig. (4-4c).

Also, when $\underline{A}(f)$ is diagonal the formulas for minimum estimation-error given by Eqs. (4-60a), (4-62a) reduce to

$$\langle \tilde{e}_o \rangle = \int_{-\infty}^{\infty} \frac{\frac{N_o}{T} \sum_{p=1}^M A_{pp}(f) |\phi_p(f)|^2}{N_o + \frac{1}{T} \sum_{q=1}^M A_{qq}(f) |\phi_q(f)|^2} df \quad (4-60b)$$

$$\langle e_o \rangle = \sum_{p=1}^M \int_{-1/2T}^{1/2T} \frac{N_o A_{pp}(f)}{N_o + A_{pp}(f)} df. \quad (4-62b)$$

Furthermore, as discussed in Subsection 2f of Chapter III, there are various CS signals of interest whose TSR correlation matrices are not only diagonal, but also constant. If we consider estimating one of these signals in additive white noise, then the solutions provided by the preceding three theorems take the particularly simple form:

$$h_o(t, \tau) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M \frac{A_o^{pp}}{N_o + A_o^{pp}} \phi_p(t-nT) \phi_p^*(\tau-nT)$$

$$\langle e_o \rangle = \frac{1}{T} \sum_{p=1}^M \frac{N_o A_o^{pp}}{N_o + A_o^{pp}}, \quad (4-63)$$

where M is infinite in Theorem (4-5), and may be finite in Theorems (4-6), (4-7).

Now if M is infinite, then $\langle e_o \rangle$ can be difficult to evaluate to within a known accuracy. With this in view, we present the following discrete version of Theorem (4-3):

THEOREM(4-8): The time-averaged value of the minimum mean-squared estimation-error which results from the optimum time-varying filter for a CS signal in additive white noise is contained within the ± 1.5 dB bounds:

$$B/2 \leq \langle e_o \rangle \leq B; \quad B = \hat{k}_{xx}(0) - \frac{1}{T} \sum_{p=1}^Q (A_o^{pp} - N_o) \quad (4-64)$$

provided that the correlation matrix of the TSR (with doubly orthonormal basis functions) for the signal is a constant diagonal matrix whose diagonal elements dominate the noise PSD for all indices less than some integer Q , and vice-versa for all indices greater than Q :

$$A_o^{pp} \begin{cases} \geq N_o, & p \leq Q \\ \leq N_o, & p \geq Q. \end{cases}$$

The proof of this theorem is simply a discrete version of that given for Theorem (4-3). Recall that Theorem (3-3) of Chapter III gives the conditions under which the correlation matrix $\underline{\Lambda}(f)$ for a CS process will be constant and diagonal, and that several examples are given. We will discuss these and other examples in the last section of this chapter.

c) Solutions based on the harmonic series representation.

i) Estimation of cyclostationary processes on an infinite interval.

As brought out in Section 3 of Chapter III, the harmonic series representation (HSR) for CS processes results directly from a specific TSR. Hence, the solution method of Theorem (4-5) must have a counterpart which employs the HSR. The solution resulting from this counterpart is stated in the following theorem:

THEOREM(4-9): The impulse-response function for the optimum time-varying filter for the general LMS estimation problem where the transmitted and received signals x, y are jointly T-CS and where y is composed of a colored component w and an uncorrelated white component with PSD λ has the representation:

$$h_o(t, \tau) = \sum_{p, q=-\infty}^{\infty} h_{pq}(t-\tau) e^{j2\pi(pt-q\tau)/T} \quad \forall t, \tau \in (-\infty, \infty) \quad (4-65)$$

where the matrix of Fourier transforms of the elements $\{h_{pq}(\cdot)\}$ is given by

$$\underline{H}(f) = \begin{cases} \underline{G}(f) [\underline{D}(f) + \lambda \underline{I}]^{-1} & \forall f \in [-1/2T, 1/2T] \\ 0 & \forall f \notin [-1/2T, 1/2T] \end{cases} \quad (4-66)$$

where \underline{D} , \underline{G} are the Fourier transforms of the bandlimited (to $[-1/2T, 1/2T]$) matrices \underline{k} , \underline{g} whose elements are given by

$$\begin{aligned} d_{pq}(\tau) &= \iint_{-\infty}^{\infty} v(\tau-\sigma) v(\gamma) k_{ww}(\sigma, \gamma) e^{-j2\pi(p\sigma-q\gamma)/T} d\sigma d\gamma \\ g_{pq}(\tau) &= \iint_{-\infty}^{\infty} v(\tau-\sigma) v(\gamma) k_{xw}(\sigma, \gamma) e^{-j2\pi(p\sigma-q\gamma)/T} d\sigma d\gamma \\ v(t) &\triangleq \frac{\sin(\pi t/T)}{\pi t}, \end{aligned} \quad (4-67)$$

and are the correlation matrices corresponding to the HSR's for x, w :

$$d_{pq}(t-s) = E\{d_p(t)d_q^*(s)\}$$

$$g_{pq}(t-s) = E\{a_p(t)a_q^*(s)\}$$

where

$$x(t) = \sum_p a_p(t)e^{j2\pi pt/T}$$

$$w(t) = \sum_p d_p(t)e^{j2\pi pt/T} \quad (4-68)$$

Furthermore, the time-averaged value of the minimum mean-squared estimation-error is, given by the formula:

$$\langle e_o \rangle = \int_{-1/2T}^{1/2T} \text{trace}[\underline{K}(f) - \underline{G}(f)[\underline{D}(f) + \lambda \underline{I}]^{-1} \underline{G}^*(f)] df \quad (4-69)$$

where $\underline{K}(f)$ is the Fourier transform of the HSR correlation matrix for x (defined as in Eq. (4-67) for w).

The proof of this theorem follows directly from Theorem (4-5), as indicated in Section 3e of Chapter III, by using Eq. (3-88) which relates the correlation matrices for TSR's and HSR's.

As discussed in Section 3e of Chapter III, a linear system whose impulse-response function has an M th order HSR, as does h_o in this theorem (with $M = \infty$), can be realized with the structure shown in Fig. (3-12) and reproduced here as Fig. (4-5). Although the number $(2M+1)$ of paths in this optimum filter will, in general, be infinite there is an interesting class of problems for which it is finite. As discussed in Subsections 3c,d,e of Chapter III, T-CS processes which are bandlimited to any interval, say $[-B, B]$ admit HSR's of order $M \leq BT$. Hence, the

optimum filters for such processes will be finite HSR structures.

The HSR structure for optimum filters has the same interpretation as the TSR structure of Fig. (4-4): The filter decomposes the received signal into jointly WSS bandlimited processes, subjects these to a matrix of optimum time-invariant filters, and then reconstructs from these filtered WSS processes the LMS estimate of the CS transmitted signal x . Thus, by employing the HSR, we convert the problem of LMS estimation of a scalar CS process into the problem of LMS estimation of a vector of jointly WSS (bandlimited) processes.

As already discussed in Section 3e of Chapter III, the matrix of time-invariant filters employed in the HSR structure have transfer functions $H_{pq}(f)$ which are bandlimited to $[-1/2T, 1/2T]$, and can therefore be exactly realized with a matrix of sampled-data filters, with transfer functions

$$\hat{H}_{pq}(f) = \sum_{n=-\infty}^{\infty} H_{pq}(f-n/T),$$

flanked at the input and output, respectively, by T -periodic impulse samplers, and ideal low-pass filters (transfer functions $V(f)$).

There is, of course, an equivalent realization of this HSR structure which employs real sinusoidal multiplier functions rather than complex exponentials (see, for example, Section 2c of Chapter III).

Paralleling the results obtained using TSR's, we have the following simplification of the formulas in Theorem (4-8): If the channel consists of nothing more than additive white noise with PSD N_0 , then the matrix solution of Eq. (4-66) reduces to

$$\underline{H}(f) = \underline{K}(f) [\underline{K}(f) + N_o \underline{I}]^{-1}, \quad (4-70)$$

which is the transfer function for a matrix Wiener filter for jointly WSS processes in additive white noise. Also, the formula of Eq. (4-59) for the time-averaged minimum mean-squared estimation-error reduces to

$$\begin{aligned} \langle e_o \rangle &= N_o \int_{-1/2T}^{1/2T} \text{trace} [\underline{K}(f) [\underline{K}(f) + N_o \underline{I}]^{-1}] df \\ &= N_o \sum_p h_{pp}(0). \end{aligned} \quad (4-71)$$

and parallels the formula (from Eq. (4-27)) for the minimum time-averaged error resulting from the optimum time-invariant filter;

$$\langle \tilde{e}_o \rangle = N_o \int_{-\infty}^{\infty} \frac{\sum_p K_{pp}(f-p/T)}{N_o + \sum_q K_{qq}(f-q/T)} df = N_o \tilde{h}_o(0). \quad (4-72)$$

(where we have used Eq. (3-89) for \tilde{K}_{xx} .)

It is interesting to note that the optimum time-invariant filter depends only on the autocorrelations of the HSR representors (the diagonal elements of the correlation matrix), and is independent of the cross-correlations. Thus, the difference between the optimum time-varying and time-invariant filters can be interpreted in terms of the fact the latter ignores the crosscorrelations, whereas the former takes full account of them.

ii) Estimation of periodic processes on a finite interval. The HSR is particularly convenient for representation of periodic processes; i.e. those whose realizations are periodic waveforms. Thus the solution of the LMS estimation problem for these processes is particularly simple when the HSR is employed. Although periodic processes are cyclostationary

processes, we can not simply apply the results of Theorem (4-9) to obtain a LMS estimate for a periodic process since the original equation for optimality (Eq. (4-12)), to which this theorem gives a solution, is not well defined for periodic processes; i.e., the integral does not exist (does not converge). This is a result of the fact that the optimum filter is not BIBO stable, and is not therefore a permissible filter (see Section 2 of this chapter). The difficulty can also be relegated to the fact that the problem of estimation of a periodic process on an infinite interval is a singular estimation problem and will result in zero mean-squared estimation-error.

Now, a more meaningful problem is that of noncausal LMS estimation of a periodic process on a finite interval. This estimation problem (the fixed-interval smoothing problem) is non-singular and is characterized by the optimality condition:

$$\int_{T_i}^{T_f} h_o(t, \sigma) k_{yy}(\tau, \sigma) d\sigma = k_{yx}(\tau, t) \quad \forall t, \tau \in [T_i, T_f] \quad (4-73)$$

which follows directly from Eq. (4-7) with $T_m = T_{rm} = |T_i - T_f| < \infty$.

For this problem, we have the following general solution:

THEOREM(4-10): The impulse-response function for the optimum fixed-interval smoother for a periodic process in additive white noise with PSD N_0 has the representation:

$$h_o(t, \tau) = \sum_{p, q} h_{pq} e^{j2\pi(pt - q\tau)/T} \quad \forall t, \tau \in [-NT/2, NT/2]$$

where the matrix of harmonic coefficients is given by

$$\underline{H} = \underline{K}[\underline{N}\underline{T}\underline{K} + \underline{N}_0\underline{I}]^{-1} \quad (4-74)$$

where the elements of the matrix \underline{K} are given by

$$k_{pq} = \frac{1}{T^2} \iint_{-T/2}^{T/2} k_{xx}(t, \tau) e^{-j2\pi(pt-q\tau)/T} dt d\tau, \quad (4-75)$$

so that \underline{K} is the correlation matrix for the HSR representors:

$$k_{pq} = E\{a_p a_q^*\}$$

where

$$x(t) = \sum_p a_p e^{j2\pi pt/T}$$

(see Sec. 3d of Chapter III). Furthermore, the time-averaged value of the minimum mean-squared estimation-error is given by the formula

$$\langle e_o \rangle = \frac{N_o}{NT} \text{trace} \left[\underline{K} \left[\underline{K} + \frac{N_o}{NT} \underline{I} \right]^{-1} \right]. \quad (4-76)$$

The proof of this theorem can be obtained simply by substituting the representations for h_o and k_{xx} , into the reduced version of Eq. (4-73):

$$\int_{-NT/2}^{NT/2} h_o(t, \sigma) k_{xx}(\tau, \sigma) d\sigma + N_o h(t, \tau) = k_{xx}(t, \tau), \quad \forall t, \tau \in [-NT/2, NT/2].$$

For comparison, we present the following general solution for the optimum time-invariant fixed-interval smoother:

THEOREM(4-11): The impulse-response function for the optimum time-invariant fixed-interval smoother for a periodic process in additive white noise with PSD N_o has the representation:

$$\tilde{h}_o(t-\tau) = \sum_q h_q e^{j2\pi q(t-\tau)/T}, \quad \forall t, \tau \in [-NT/2, NT/2]$$

where

$$h_q = \frac{k_q}{NTk_{-q} + N_o} \quad (4-77)$$

and where $\{k_q\}$ are the Fourier series coefficients for the periodic autocorrelation function for the stationarized version of the periodic process, and are identical to the diagonal elements of the HSR correlation matrix:

$$k_q = k_{qq}.$$

Furthermore, the minimum value of the time-averaged estimation error is given by the formula:

$$\langle \tilde{e}_0 \rangle = \frac{N_o}{NT} \sum_q \frac{k_q}{k_{-q} + N_o/NT}. \quad (4-78)$$

The proof of this theorem, like that for the preceding theorem, can be obtained simply by substituting Eq. (4-77) for \tilde{h}_0 , and the Fourier series representation for \tilde{k}_{xx} , into the optimality condition:

$$\int_{-NT/2}^{NT/2} \tilde{h}_0(t-\sigma) \tilde{k}_{xx}(\tau-\sigma) d\sigma + N_o \tilde{h}_0(t-\tau) = \tilde{k}_{xx}(t-\tau) \quad \forall t, \tau \in [-NT/2, NT/2]. \quad (4-79)$$

Note that as the length of the observation interval approaches infinity, the estimation errors of Theorems (4-10), (4-11) both approach zero, as expected from the preliminary discussion.

The optimum smoothers of Theorems (4-10), (4-11) can be realized with the structure shown in Fig. (4-6), where M is the highest non-zero harmonic in the signal being smoothed. The numbers $\{\hat{a}_n\}$ at the outputs of the matrix of attenuators are the LMS linear estimates of the Fourier series coefficients of the periodic signal x . For the time-invariant smoother of Theorem (4-11), the matrix of attenuators is diagonal ($h_{pp} = h_n$) so that the overall structure becomes a parallel connection of $2M+1$ paths.

From this it can be seen that the difference between the optimum time-varying and time-invariant smoothers is that the latter ignores the crosscorrelation between various harmonics.

d) Solutions based on the Fourier series representation for correlation functions. We begin with a theorem:

THEOREM(4-12): The impulse-response function for the optimum filter for the general LMS estimation problem where the transmitted and received signals x, y are jointly T-CS and where y is composed of a colored component w and an uncorrelated white component with PSD λ has the representation:

$$h_o(t, \tau) = \sum_{n=-\infty}^{\infty} h_n(t-\tau) e^{j\pi n(t+\tau)/T} \quad \forall t, \tau \in (-\infty, \infty) \quad (4-80)$$

where the Fourier transforms $\{H_n\}$ of the time-invariant impulse-response function $\{h_n\}$ are the solutions of the linear algebraic equations:

$$\sum_{m=-\infty}^{\infty} C_{n-m}(f-(n+m)/2T) H_m(f-m/2T) + \lambda H_n(f-n/2T) = G_n(f-n/2T) \quad \forall n \quad (4-81)$$

where the $\{C_n\}$ and $\{G_n\}$ are the Fourier transforms of the coefficient-functions for the PSR's for the correlation functions k_{ww}, k_{xw} :

$$\begin{aligned} c_n(t) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{ww}(s+t/2, s-t/2) e^{-j2\pi ns/T} ds \\ g_n(t) &= \frac{1}{T} \int_{-T/2}^{T/2} k_{xw}(s+t/2, s-t/2) e^{-j2\pi ns/T} ds \end{aligned} \quad (4-82)$$

(see Section 4a of Chapter III).

The proof of this theorem is outlined in terms of Eqs. (3-117) - (3-121) in Section 4c of Chapter III.

There does not appear to be any way to obtain a general closed-form solution to this infinite set of equations; however, if the functions

$\{C_n\}$, $\{G_n\}$ satisfy certain bandlimiting constraints, then solutions are possible, as discussed in Section 4c of Chapter III. For example:

THEOREM(4-13): The impulse-response functions for the optimum filter for a T-CS process x in additive white noise with PSD N_0 has the representation:

$$h_0(t, \tau) = \sum_{n=-\infty}^{\infty} h_n(t-\tau) e^{j\pi n(t+\tau)/T} \quad (4-83)$$

where the Fourier transforms $\{H_n\}$ of the functions $\{h_n\}$ are given by the formulas:

$$H_n(f) = \frac{C_n(f) [1 - H_0(f+n/2T)]}{N_0 + C_0(f-n/T)},$$

$$H_0(f) = \frac{R(f)}{N_0 + R(f)}, \quad (4-84)$$

where

$$R(f) = C_0(f) - \sum_{n \neq 0} \frac{|C_n(f-n/2T)|^2}{N_0 + C_0(f-n/T)}, \quad (4-85)$$

provided that

$$C_n(f) = 0, \quad |f| \geq 1/4T, \quad \forall n \neq 0, \quad (4-86)$$

where $\{C_n\}$ are the Fourier transforms of the FSR coefficient functions:

$$c_n(t) = \frac{1}{T} \int_{-T/2}^{T/2} k_{xx}(s+t/2, s-t/2) e^{j2\pi ns/T} ds. \quad (4-87)$$

Furthermore, the time-averaged value of the minimum mean-squared estimation-error is given by the formula:

$$\langle e_0 \rangle = \int_{-\infty}^{\infty} \frac{N_0 R(f)}{N_0 + R(f)} df. \quad (4-88)$$

Proof: Using the condition of Eq. (4-86) in Eq. (4-81) of Theorem (4-12), where $G_n = C_n$, $\lambda = N_0$, yields the equations:

$$\sum_{m \neq 0} C_{-m}(f-m/2T)H_m(f-m/2T) + (C_0(f)+N_0)H_0(f) = C_0(f), \quad (4-89)$$

$$(C_0(f-n/T)+N_0)H_n(f-n/2T) + C_n(f-n/2T)H_0(f) = C_n(f-n/2T) \quad \forall n \neq 0. \quad (4-90)$$

Now, Eqs. (4-84), (4-85) follow directly from Eqs. (4-89), (4-90).

Furthermore, the formula of Eq. (4-88) results from substituting Eqs. (4-83), (4-84) into the formula of Eq. (4-32).

QED

Note that if k_{xx} has an M^{th} order FSR satisfying the constraint of Eq. (4-86), then so too does the solution h provided by Theorem (4-13). As discussed in Section 4c of Chapter III, a linear system whose impulse-response function has an M^{th} order FSR can be realized with the $(2M+1)$ -path structure shown in Fig. (3-14a), or with the simplified structure shown in Fig. (3-14b) and reproduced here as Fig. (4-7), where the impulse-response functions for the path filters with transfer functions G_{n1} , G_{n2} are related to the original coefficient-functions as follows:

$$\begin{aligned} g_{n1}(t) &= 2\text{Re}\{h_n(t)\}\cos(\pi nt/T) - 2\text{Im}\{h_n(t)\}\sin(\pi nt/T) \\ g_{n2}(t) &= -2\text{Re}\{h_n(t)\}\sin(\pi nt/T) - 2\text{Im}\{h_n(t)\}\cos(\pi nt/T). \end{aligned} \quad (4-91)$$

Note that, as mentioned in subsection 4c of Chapter III, the solution method of Theorem (4-13) can be extended to the case where the Fourier coefficient functions satisfy the less restrictive constraint

$$C_n(f) = 0, \quad |f| \geq M/4T, \quad \forall n \neq 0.$$

However, the method then requires inversion of a $(2M-1) \times (2M-1)$ matrix of functions, and does not appear to offer any advantage over the solution method based on the HSR, which is considerably less cumbersome.

e) Solutions based on inverse-operator and feedback-system methods.

i) Introduction. The basic integral equation derived in the second section of this chapter:

$$\int_{-\infty}^{\infty} h_o(t,s)k_{yy}(s,\tau)ds = k_{xy}(t,\tau) \quad \forall t,\tau \in (-\infty,\infty), \quad (4-92)$$

which implicitly specifies the impulse-response function for the optimum filter, can be interpreted as an operator equation. Specifically, if we denote the linear integral transformations (operators) with kernels $k_{yy}(\cdot,\cdot)$, $h_o(\cdot,\cdot)$, $k_{xy}(\cdot,\cdot)$ as \underline{K}_{yy} , \underline{H}_o , \underline{K}_{xy} , then the operator counterpart of this integral equation can be written as

$$\underline{H}_o \cdot \underline{K}_{yy} = \underline{K}_{xy}. \quad (4-93)$$

This equation has the formal solution

$$\underline{H}_o = \underline{K}_{xy} \cdot \underline{K}_{yy}^{-1}, \quad (4-94)$$

where \underline{K}_{yy}^{-1} is the inverse operator corresponding to \underline{K}_{yy} :

$$\underline{K}_{yy} \cdot \underline{K}_{yy}^{-1} = \underline{I}, \quad (4-95)$$

where \underline{I} is the identity operator.

Now, if we denote the kernel for this inverse operator as $k_{yy}^{-1}(\cdot,\cdot)$, then the impulse-response function for the optimum filter is formally

given by

$$h_0(t, \tau) = \int_{-\infty}^{\infty} k_{xy}(t, \sigma) k_{yy}^{-1}(\sigma, \tau) d\sigma. \quad (4-96)$$

Furthermore, this optimum filter can be realized, as shown in Fig. (4-8), as the cascade connection of the two linear systems which are characterized by the operators \underline{K}_{yy}^{-1} , \underline{K}_{xy} , and have impulse-response functions $k_{yy}^{-1}(\cdot, \cdot)$, $k_{xy}(\cdot, \cdot)$.

There are a number of interesting LMS estimation problems for T-CS processes where the kernels of the operators \underline{K}_{yy} , \underline{K}_{xy} can be decomposed in such a manner that the system diagram of Fig. (4-8) can be manipulated into the feedback configuration shown in Fig. (4-9), where \underline{K}_0 is a known time-invariant system, $p(\cdot)$ is a known T-periodic function, and \underline{K}_1 , \underline{K}_2 are known periodically (T) time-varying systems. Furthermore, there is a non-trivial set of circumstances under which the impulse-response function h_0 for this system can be obtained explicitly:

Namely,

$$\begin{aligned} K_0(f) &= 0, & ||f| - M/2T| &\geq N/2T \\ p_n &= 0, & ||n| - M| < N, &\forall n \neq 0 \end{aligned} \quad (4-97)$$

where N , M are nonnegative integers, $K_0(\cdot)$ is the transfer function for \underline{K}_0 , and $\{p_n\}$ are the Fourier coefficients for $p(\cdot)$.

For $M \geq 1$, \underline{K}_0 is a bandpass filter, and for $M = 0$, \underline{K}_0 is a lowpass filter. For $N = 1$, $M = 0$, $p(\cdot)$ is unconstrained, and for $N > 1$, $M = 0$, $p(\cdot)$ contains only frequency-components which are too high to be passed by the filter \underline{K}_0 , and for $N = 1$, $M > 1$, $p(\cdot)$ contains frequency components

which are too low to be passed by K_0 . Now, when the conditions of Eq. (4-97) are satisfied, the system of Fig. (4-9) is unaffected by the replacement of the periodic multiplier in the feedback path with a constant attenuator whose attenuation is p_0 , the average value of $p(\cdot)$. Thus, the entire feedback loop is time-invariant, and has transfer function K_3 given by

$$K_3(f) = \frac{(1/p_0)K_0(f)}{(1/p_0) + K_0(f)}, \quad (4-98)$$

so that the impulse-response function for the entire system of Fig. (4-9) is

$$h_0(t, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_2(t, \sigma) k_3(\sigma, s) k_1(s, \tau) ds d\sigma, \quad (4-99)$$

where k_3 is the inverse Fourier transform of K_3 .

We refer to this method of solving for optimum time-varying filters as the inverse-operator-feedback-system (IOFS) method. We now present four classes of practical optimum-filtering problems for which the general solution of Eqs. (4-98), (4-99) is valid. In the next section, we illustrate these classes of problems with specific problems.

ii) Examples:

1) WSS signal subjected to time-invariant dispersion, periodic attenuation, and additive WSS white noise: For the class of optimum-filtering problems where a WSS signal is transmitted through a channel consisting of a time-invariant dispersion with impulse-response function g , T -periodic attenuation q , and additive WSS white noise with PSD N_0 , the correlation functions in Eq. (4-92) take the form:

$$\begin{aligned}
 k_{yy}(t,s) &= \iint_{-\infty}^{\infty} k_{xx}(\sigma-\gamma)g(t-\sigma)g(s-\gamma)d\sigma d\gamma q(t)q(s) + N_0\delta(t-s) \\
 k_{xy}(t,s) &= \int_{-\infty}^{\infty} k_{xx}(t-\sigma)g(s-\sigma)d\sigma q(s),
 \end{aligned}
 \tag{4-100}$$

so that the systems, with these functions as impulse responses, take the forms shown in Fig. (4-10). Now, using elementary algebra for linear systems [36], it is easily shown that the system of Fig. (4-8) is equivalent to that of Fig. (4-9), where the impulse-response functions are given by

$$\begin{aligned}
 k_1(t,s) &= q(t)\delta(t-s) \\
 k_2(t) &= F^{-1}\{E_{xx}(f)|G(f)|^2\}/N_0 \\
 k_3(t) &= F^{-1}\{1/G(f)\},
 \end{aligned}
 \tag{4-101}$$

and where $p(t) = \sigma^2(t)$. Now if the channel dispersion or the signal PSD are bandpass, and the periodic fluctuations in attenuation are relatively low-frequency, or if the dispersion or PSD are low-pass, and the fluctuations in attenuation are relatively high-frequency, (so that the conditions of Eq. (4-97) are satisfied), then the general solutions of Eqs. (4-98), (4-99), for the optimum filter, are valid.

(2) CS AS signals in additive WSS white noise: For the class of optimum-filtering problems where a CS signal, which is the product of a deterministic T-periodic function q , and a WSS process z , is transmitted through a channel consisting only of additive WSS white noise with PSD N_0 , the correlation functions in Eq. (4-91) take the forms

$$\begin{aligned}
 k_{yy}(t,s) &= q(t)k_{zz}(t-s)q(s) + N_0\delta(t-s) \\
 k_{xy}(t,s) &= q(t)k_{zz}(t-s)q(s),
 \end{aligned}
 \tag{4-102}$$

so that the systems, with these functions as impulse-responses, take the forms shown in Fig. (4-11). Using the algebra for linear systems, it is easily shown that the system of Fig. (4-8) is equivalent to that in Fig. (4-9), where the impulse-response functions are given by

$$k_1(t,s) = q(t)\delta(t-s)$$

$$k_0(\tau) = k_{zz}(\tau)/N_0$$

$$k_2(t,s) = q(t)\delta(t-s), \quad (4-103)$$

and where $p(t) = q^2(t)$.

Now if the "baseband" signal z is lowpass and the periodic fluctuations in the "carrier" are relatively high-frequency, or if the process z is bandpass, and the multiplicative factor q is relatively slow-varying (so that the conditions of Eq. (4-97) are satisfied), then the general solution of Eqs. (4-98), (4-99) for the optimum filter, are valid.

(5) WSS signal in additive CS white noise: For the class of optimum-filtering problems where a WSS signal is transmitted through a channel consisting of only additive CS white noise with autocorrelation function $q^2(\tau)\delta(t-s)$, the correlation functions in Eq. (4-92) take the forms:

$$k_{yy}(t,s) = k_{xx}(t-s) + q^2(t)\delta(t-s)$$

$$k_{xy}(t,s) = k_{xx}(t-s), \quad (4-104)$$

so that the systems, with these functions as impulse responses, take the forms shown in Fig. (4-12). Using the algebra of linear systems, it is

again easily shown that the system of Fig. (4-3) is equivalent to that in Fig. (4-9), where the impulse-response functions are given by

$$\begin{aligned} k_1(t,s) &= q^{-2}(t)\delta(t-s) \\ k_0(\tau) &= k_{xx}(\tau) \\ k_2(t,s) &= \delta(t-s), \end{aligned} \quad (4-105)$$

and where $p(t) = q^{-2}(t)$.

Now if the periodic fluctuations in the intensity of the additive white noise are either low-frequency compared to the bandpass signal, or are high-frequency compared to the lowpass signal (so that Eq. (4-97) is satisfied), then the general solution of Eqs. (4-98), (4-99), for the optimum filter, are valid.

(4) CS signal (obtained from periodically time-scaled WSS process) in additive WSS white noise: For the class of optimum-filtering problems where a CS signal, which is obtained from a WSS process z via a periodic time-scale transformation $g(t) = t + q(t)$ (as discussed in Section 2c of Chapter II), is transmitted through a channel consisting only of additive WSS white noise with PSD N_0 , the correlation functions in Eq. (4-92) take the forms:

$$\begin{aligned} k_{yy}(t,s) &= k_{zz}(g(t)-g(s)) + N_0\delta(t-s) \\ k_{xy}(t,s) &= k_{zz}(g(t)-g(s)), \end{aligned} \quad (4-106)$$

so that the systems, with these functions as impulse responses, take the forms shown in Fig. (4-13). Using the algebra of linear systems, it is easily shown that the system of Fig. (4-8) is equivalent to that of Fig. (4-9), where the impulse-response functions are given by

$$k_1(t,s) = \delta(t-g(s))$$

$$k_0(\tau) = k_{zz}(\tau)/N_0$$

$$k_2(t,s) = \delta(g(t)-s) \quad (4-107)$$

and where $p(t) = 1/g'(t)$, where the prime indicates differentiation.

Now if the periodic fluctuations in the time-scale transformation (i.e., in $1/(1+q'(t))$) are either low-frequency compared to the bandpass WSS process z , or are high-frequency compared to the lowpass WSS process z (so that Eq. (4-97) is satisfied), then the general solution of Eqs. (4-89), (4-99) are valid.

6. Examples of Optimum Filters and Improvements in Performance

a) Introduction. In this section we present a number of specific examples of solutions for optimum filters for cyclostationary processes, and of evaluations of the performance index I (defined in Eq. (4-28)), which indicates the degree of improvement in the performance of the optimum time-varying filter over that of the optimum time-invariant filter. We classify the examples according to the method of solution employed, so that the examples in each of the subsections 6b-6e correspond to the methods presented in each of the subsections 5b-5e.

For simplicity, we consider only those estimation problems where the channel consists of nothing more than additive white noise with PSD N_0 (unless otherwise stated.)

b) Examples employing translation series representations.

i) Frequency-shift-keyed signals. As a specific example of a synchronous M -ary signal, we consider the frequency-shift-keyed signal (Model (12) of Chapter II) which takes the form

$$x(t) = \sum_{n=-\infty}^{\infty} w(t-nT) \cos(2\pi a_n f_0 t), \quad (4-108)$$

where the frequency-factors $\{a_n\}$ are statistically independent random variables which take on each of the values in the M -ary alphabet $\{1, 2, \dots, M\}$ with probability $1/M$, and where w is a sinc-pulse:

$$w(t) = \sqrt{2B} \frac{\sin(\pi B t)}{(\pi B t)}, \quad (4-109)$$

where the single-sided bandwidth $B \leq f_0$, and where $f_0 T$ is some positive integer.

As shown in Secs. 2e, 2f of Chapter II, x admits an M^{th} order TSR with the orthonormal basis functions

$$\phi_p(t) = w(t)\cos(2\pi f_0 t) \quad (4-110)$$

and correlation matrix (for the centered process $x - m_x$)

$$A_{pq}(f) = \frac{1}{M} \delta_{pq} - \frac{1}{M^2} \quad (4-111)$$

Now, from Theorem (4-6), the optimum (minimum error-variance)²⁸ time-varying filter for this FSK signal in white noise has impulse-response function given by Eq. (4-56), and has the realization shown in Fig. (4-4), where the transfer function for the matrix of sampled-data filters is given by Eq. (4-61a). The structure of this optimum filter has an interesting interpretation. The bank of input filters are matched to each of the M possible pulses that occur in the signal. The periodically sampled outputs of these matched filters are passed through a discrete matrix Wiener filter, the outputs of which can be shown to be the LMS estimates of the a-posteriori probabilities; i.e., the p^{th} output at time nT is the LMS estimate of the probability that the n^{th} symbol transmitted was the p^{th} letter of the M -ary alphabet, $a_n = p$, given that y was received [62]. These probability estimates are used to weight the pulses which are regenerated at the output of the filter. See Fig. (4-14).

The time-averaged minimum estimation-error is given by Eq. (4-62a), and reduces to the expression

²⁸ See the last paragraph at the end of section 2 of this chapter for a discussion of the relationship between minimum-mean-squared-error and minimum-error-variance.

$$\begin{aligned}
\langle e_o \rangle &= \frac{N_o}{T} \text{trace } [A_o + N_o I]^{-1} \\
&= \frac{1}{T} \frac{N_o (M-1)}{MN_o + 1}
\end{aligned} \tag{4-112}$$

(using Theorem (3-4)).

The optimum time-invariant filter has transfer function given by Eq. (4-59), and the resultant minimum time-averaged estimation-error is given by Eq. (4-60a):

$$\langle \tilde{e}_o \rangle = \sum_{p=1}^M \int_{-\infty}^{\infty} \frac{\frac{N_o}{T} A_o^{pp} |\phi_p(f)|^2}{N_o + \frac{1}{T} A_o^{pp} |\phi_p(f)|^2} df, \tag{4-113}$$

where we have used the fact that $\phi_p(f) \phi_q^*(f) = |\phi_p(f)|^2 \delta_{pq}$.

Now, since the $\{\phi_p\}$ are rectangular with height $1/\sqrt{2B}$ and double-sided bandwidth $2B$, then

$$\langle \tilde{e}_o \rangle = 2BMN_o \frac{(M-1)/2BTM}{(M-1)/2BTM + MN_o}. \tag{4-114}$$

Using an effective noise-bandwidth of $2BM$ to obtain a noise variance of $2BMN_o$, and using the time-averaged signal-variance $(M-1)/MT$, we can express the performance index of Eq. (4-28) in terms of the signal-to-noise ratio (of variances) ρ :

$$I \Delta \frac{\langle \tilde{e}_o \rangle}{\langle e_o \rangle} = \frac{1 + \rho 2BTM/(M-1)}{1 + \rho}. \tag{4-115}$$

The lowest practical bandwidth of the envelope w corresponds to transmitting pulses at the Nyquist rate, and is $B = 1/T$. This results in a minimum improvement factor (value of performance index) of 6 dB for binary FSK ($M = 2$) in low noise. As the bandwidth is increased in order to decrease intersymbol interference, the low noise improvement factor

increases in direct proportion, which means that the time-varying filter becomes more attractive, relative to the time-invariant filter, as the quality of transmission increases.

It is interesting to note that for signal-to-noise ratios as low as unity, there can still be a substantial improvement:

$$I = 1/2 + BTM/(M-1) \quad \text{for } \rho = 1. \quad (4-116)$$

ii) Time-division-multiplexed PAM signals. An interesting (and perhaps more practical) alternative to the time-division-multiplexing schemes TDM_1 , TDM_2 discussed in Chapter II (Models (3), (5)) is the TDM of PAM scheme, which was briefly mentioned in Sec. 2f of Chapter III. For this scheme, we consider M mutually uncorrelated WSS component signals $\{x_p\}$ with autocorrelation functions $\{k_p\}$ and PSD's $\{K_p\}$ which are bandlimited to the interval $[-1/2T, 1/2T]$. The signals are periodically sampled every T seconds to amplitude-modulate orthonormal pulses $\{\phi(t-pT/M); p = 1, 2, \dots, M\}$, so that the composite TDM-PAM signal takes the form

$$x(t) = \sqrt{T/M} \sum_{p=1}^M \sum_{n=-\infty}^{\infty} x_p(nT) \phi(t-pT/M-nT). \quad (4-117)$$

Clearly x admits an M^{th} order TSR, and since the $\{x_p\}$ are uncorrelated, then the correlation matrix is diagonal with on-diagonal elements

$$A_{pp}(f) = \frac{T}{M} \sum_n k_p(nT) e^{-j2\pi nTf} = \frac{1}{M} \sum_m K_p(f+m/T). \quad (4-118)$$

Now, from Theorem (4-6), the optimum time-varying filter for this TDM signal in white noise has impulse-response function given by Eq. (4-56), and has the realization shown in Fig. (4-4), where the transfer function for the diagonal matrix of sampled-data filters is given by Eq. (4-61b).

The structure of this optimum filter has an interesting interpretation. Since the matrix of sampled-data filters \underline{H} is diagonal, then the overall structure is simply the parallel connection of M independent paths as shown in Fig. (4-4c). The p^{th} path consists of a discrete Wiener filter, for the p^{th} sequence of amplitudes, flanked at the input and output, respectively, by a pulse-amplitude demodulator and a pulse-amplitude modulator. Hence, the optimum time-varying filter demodulates the signal, optimally filters it at "baseband", and then remodulates it. See Fig. (4-15).

The time-averaged minimum estimation-error is given by Eq. (4-62b) which reduces to the expression

$$\langle \tilde{e}_o \rangle = N_o \sum_{p=1}^M \int_{-1/2T}^{1/2T} \frac{K_p(f)}{MN_o + K_p(f)} df. \quad (4-119)$$

The optimum time-invariant filter has transfer function given by Eq. (4-59), and the resultant minimum time-averaged error is given by Eq. (4-60b). If we assume that the pulse ϕ is the interpolating sinc-pulse whose Fourier transform is ideal low-pass with double-sided bandwidth M/T and height $\sqrt{T/M}$, then Eq. (4-60b) reduces to

$$\langle \tilde{e}_o \rangle = N_o \int_{-1/2T}^{1/2T} \frac{\sum_{p=1}^M K_p(f)}{MN_o + \frac{1}{M} \sum_{q=1}^M K_q(f)} df. \quad (4-120)$$

Now, if the M processes $\{x_p\}$ are identical, then the TDM process degenerates to a WSS process, and $I = 1$: there is no improvement. However, if the individual processes are different, there can be substantial improvement. For example, if the processes have disjoint PSD's, then Eq. (4-120) reduces to

$$\langle \tilde{e}_o \rangle = N_o \sum_{p=1}^M \int_{-1/2T}^{1/2T} \frac{K_p(f)}{MN_o + \frac{1}{M} K_p(f)} df, \quad (4-121)$$

and there is a low-noise improvement factor of $I = M$. Similarly, if all but one (say the q^{th}) of the M baseband signals are zero, then the formulas of Eqs. (4-119), (4-120) for estimation-error become :

$$\langle e_o \rangle = N_o \int_{-1/2T}^{1/2T} \frac{K_q(f)}{MN_o + K_q(f)} df \quad (4-122)$$

$$\langle \tilde{e}_o \rangle = N_o \int_{-1/2T}^{1/2T} \frac{K_q(f)}{MN_o + \frac{1}{M} K_q(f)} df, \quad (4-123)$$

and there is again a low-noise improvement factor of $I = M$.

iii) Frequency-division-multiplexed signals. A multiplexing scheme which is in some ways a frequency-time dual of TDM schemes is the frequency-division-multiplexing (FDM) scheme discussed in Chapter II (Model (4)). We consider here FDM signals of the form

$$x(t) = \sum_{p=1}^M x_p(t) \cos(\omega_o t + 2\pi p t / T_o) \quad (4-124)$$

where ω_o is an integer-multiple of $2\pi/T$, and T is an integer-multiple of T_o , and where the M component signals $\{x_p\}$ with correlation functions $\{k_p\}$ are assumed WSS, uncorrelated, and bandlimited to $[-1/2T, 1/2T]$.

From the sampling theorem (Theorem (2-2)), the component signals admit the representations

$$x_p(t) = \sum_n x_p(nT) \frac{\sin \pi(t-nT)/T}{\pi(t-nT)/T} \quad (4-125)$$

so that the composite FDM signal admits the M^{th} order TSR

$$x(t) = \sqrt{T/2} \sum_{p=1}^M \sum_{n=-\infty}^{\infty} x_p(nT) \phi_p(t-nT), \quad (4-126)$$

where the basis functions are given by

$$\phi_p(t) = \sqrt{2/T} \cos(\omega_0 t + 2\pi p t/T_0) \frac{\sin(\pi t/T)}{(\pi t/T)}. \quad (4-127)$$

These basis functions are doubly orthonormal, and the TSR correlation matrix $\underline{A}(f)$ is diagonal with on-diagonal elements given by

$$A_{pp}(f) = \frac{T}{2} \sum_n k_p(nT) e^{-j2\pi n T f} = \frac{1}{2} \sum_m K_p(f+m/T). \quad (4-128)$$

The duality with the TDM-PAM scheme is now evident since the correlation matrices are identical (except for a scale factor), and the basis functions are time-frequency duals (if the pulse-shape used in the TDM scheme is the interpolating sinc-pulse) in the sense that the M FDM basis functions are frequency-translated sinc-pulses, and the M TDM basis functions are time-translated sinc-pulses.

Now, from Theorem (4-6), the optimum impulse-response function is given by Eq. (4-56), and has the realization shown in Fig. (4-4c), where the transfer functions for the diagonal matrix of sampled-data filters are given by Eq. (4-61b). As expected, the structure of this optimum filter parallels that for the TDM-PAM signal. As shown in Fig. (4-16), it consists of M independent paths. The p^{th} path contains an ideal bandpass input filter (which passes only the p^{th} component signal x_p) followed by a periodic sampler, and a discrete Wiener filter, and an ideal bandpass output filter, which reconstructs a smooth signal from the filtered samples.

The time-averaged minimum estimation-error is given by Eq. (4-62b) which reduces to

$$\langle e_o \rangle = N_o \sum_{p=1}^M \int_{-1/2T}^{1/2T} \frac{K_p(f)}{2N_o + K_p(f)} df. \quad (4-129)$$

The optimum time-invariant filter has transfer function given by Eq. (4-59), and the resultant minimum estimation-error is given by Eq. (4-60b) which reduces to

$$\langle \tilde{e}_o \rangle = N_o \sum_{p=1}^M \int_{-1/2T}^{1/2T} \frac{K_p(f)}{2N_o + \frac{1}{2} K_p(f)} df. \quad (4-130)$$

From these last two formulas, it is clear that for low-noise conditions, the optimum time-varying filter provides a 3 dB improvement over the optimum time-invariant filter, regardless of the specific shapes or differences amongst the M PSD's, and regardless of the number (M) of signals multiplexed; i.e., there is an independent 3 dB improvement for every component signal x_p .

We will consider the FDM signal again in the next section where we employ the HSR to gain some additional insight into the structure of the optimum filter.

iv) Pulse-amplitude modulated signals. The PAM signal (Model (2) of Chapter II),

$$x(t) = \frac{1}{\phi(0)} \sum_{n=-\infty}^{\infty} a_n \phi(t-nT) \quad (4-131)$$

is one of the simplest, and yet one of the richest examples of a CS process. Every PAM process admits a first order TSR--in fact, as shown in Sec. 2 of Chapter III, all T-CS processes (in H_{CS}) can be represented in terms of T-CS PAM processes with full duty-cycle rectangular pulses, and such PAM processes have constant means and variances. Furthermore, as shown in Sec. 3 of Chapter III, all T-CS processes can also be represented in terms of WSS PAM processes with bandlimited ($\text{to}[-1/2T, 1/2T]$)

For now,

pulses. We now, use these two types of PAM processes (those where either the pulse ϕ or its Fourier transform $\hat{\phi}$ are rectangular) to demonstrate three interesting results concerning optimum filtering of CS processes.

Before proceeding, however, we state here the simplified versions of Eqs. (4-58), (4-60a) for estimation-errors of arbitrary PAM signals in additive white noise:

$$\begin{aligned} \langle e_o \rangle &= N_o \int_{-1/2T}^{1/2T} \frac{\frac{1}{T} |\hat{\phi}(f)|^2 A(f)}{N_o + \frac{1}{T} |\hat{\phi}(f)|^2 A(f)} df \\ \langle \tilde{e}_o \rangle &= N_o \int_{-\infty}^{\infty} \frac{\frac{1}{T} |\phi(f)|^2 A(f)}{N_o + \frac{1}{T} |\phi(f)|^2 A(f)} df \end{aligned}$$

where

$$\begin{aligned} A(f) &= \frac{1}{\phi^2(0)} \sum_n E\{a_{n+m} a_m\} e^{-j2\pi n T f} \\ |\hat{\phi}(f)|^2 &= \sum_n |\phi(f+n/T)|^2. \end{aligned} \quad (4-132)$$

(1) Intuition can sometimes be misleading. In particular, one might intuitively expect that a cyclostationary process with constant mean and variance would be "almost" stationary, and that no substantial improvement would result from using periodic filters for estimation of such processes. As a counterexample, consider PAM with full duty-cycle unit-energy rectangular pulses. This process admits a 1st order TSR, and has constant mean and variance. If the PSD $K(f)$ of the WSS process, whose samples $\{a_n\}$ are the pulse-amplitudes, is bandlimited to $[-1/2T, 1/2T]$, then from Eq. (4-132), we obtain the estimation-errors:

$$\begin{aligned}
\langle e_o \rangle &= \int_{-1/2T}^{1/2T} \frac{N_o K(f)}{N_o + K(f)} df \\
\langle \tilde{e}_o \rangle &= \sum_{n=-\infty}^{\infty} \int_{-1/2T}^{1/2T} \frac{N_o K(f) [\sin \pi(Tf-n)/\pi(Tf-n)]^2}{N_o + K(f) [\sin \pi(Tf-n)/\pi(Tf-n)]^2} df \quad (4-133)
\end{aligned}$$

Now, for low noise, $\langle e_o \rangle \approx N_o/T$, and for any number J there is an N_o small enough to guarantee that $\langle \tilde{e}_o \rangle > JN_o/T$. For example, an improvement factor $I > J = 10$ results with a signal-to-noise ratio of $\sigma^2 T/N_o = 10^4$, where we have chosen $K(f) = \sigma^2 T$, $f \in [-1/2T, 1/2T]$.

(2) Another situation where intuition might lead one astray is in the case of high noise. One might expect that signal-to-noise ratios as low as unity would degrade time-varying filter performance to the point where improvement over that of the optimum time-invariant filter would be negligible. As a counterexample, we have the results for PSK. In addition, we again consider PAM but this time with unit-energy ideal lowpass pulses with bandwidth M/T . From Eq. (4-132), with $K(f)$ constant for $f \in [-1/2T, 1/2T]$, these equations reduce

$$\langle e_o \rangle = \frac{\sigma_s^2 \sigma_n^2 / M}{\sigma_s^2 + \sigma_n^2 / M}, \quad \langle \tilde{e}_o \rangle = \frac{\sigma_s^2 \sigma_n^2}{\sigma_s^2 + \sigma_n^2} \quad (4-134)$$

where σ_s^2 is the time-averaged signal variance, and $\sigma_n^2 = N_o M/T$ is the noise variance. Notice that the effective noise variance for the time-varying filter is reduced by the factor $1/M$. If we fix the signal-to-noise ratio at unity ($\sigma_s^2 = \sigma_n^2$), then the improvement factor is $I = (M+1)/2$. If M is substantially greater than unity ($M = 1$ means transmission at the Nyquist rate and renders the PAM signal WSS), then the time-varying filter will provide substantial improvement in performance over that of the time-invariant filter for this high noise situation.

(3) In contrast to the two above results, the following result is very much in agreement with intuition. By considering PAM with less than full duty-cycle rectangular pulses, we find that improvement increases as the signal becomes "more" cyclostationary. Specifically, consider a unit-energy rectangular pulse of width T/M , and uncorrelated pulse-amplitudes, so that the Fourier transform of the pulse is given by

$$\phi(f) = \sqrt{T/M} \frac{\sin(\pi f T/M)}{(\pi f T/M)},$$

and the one-dimensional correlation matrix is given by

$$\underline{A}(f) = T\sigma_s^2$$

where σ_s^2 is the time-averaged signal variance. Now, using Eq. (4-132)

we obtain the estimation-errors:

$$\langle e_o \rangle = \frac{\sigma_s^2 N_o / T}{\sigma_s^2 + N_o / T} \quad (4-135)$$

$$\begin{aligned} \langle \tilde{e}_o \rangle &= \int_{-\infty}^{\infty} \frac{N_o \sigma_s^2 |\phi(f)|^2}{N_o + \sigma_s^2 |\phi(f)|^2} df \\ &> \int_{-\infty}^{\infty} \frac{N_o \sigma_s^2 |\phi(f)|^2}{N_o + \sigma_s^2 T/M} df \\ &= \frac{\sigma_s^2 N_o / T}{\sigma_s^2 / M + N_o / T} \end{aligned} \quad (4-136)$$

Hence, for low noise, we have an improvement factor of $I > M$. Clearly, as M is increased, the rectangular pulses become narrower so that the cyclostationarity is enhanced, and improvement is increased.

v) PAM signals through random channels. We now consider the transmission of PAM signals through random channels consisting of either a constant, but random attenuation and additive white noise, or of a constant, but random, delay and additive white noise. We assume that the pulse-translates $\{\phi(t-nT)\}$ are orthonormal.

(1) Random delay: We assume that the channel consists of a constant, but random, delay θ with PDF $p_\theta(\cdot)$, and additive white noise with PSD N_0 , and we assume that the phase of the received signal has been determined by a "tracking receiver" (see Subsection 7b of Chapter II and [40]). We consider the problem of estimating the transmitted signal x with its original phase, assuming this phase information to be of some importance. For this LMS estimation problem, we have the correlation functions

$$\begin{aligned} k_{yy}(t,s) &= k_{xx}(t,s) + N_0 \delta(t-s) \\ &= \sum_{n,m} A_{n-m} \phi(t-nT) \phi(s-mT) + N_0 \delta(t-s) \\ k_{xy}(t,s) &= \int_{-\infty}^{\infty} k_{xx}(t+\sigma, s) p_\theta(\sigma) d\sigma \\ &= \sum_{n,m} A_{n-m} \bar{\phi}(t-nT) \phi(s-mT) \end{aligned}$$

where

$$\bar{\phi}(t) = \int_{-\infty}^{\infty} \phi(t+\sigma) p_\theta(\sigma) d\sigma. \quad (4-137)$$

By employing the solution method of Theorems (4-5), (4-6), (4-7) we easily obtain the solution for the impulse-response function for the optimum filter:

$$h_o(t, \tau) = \sum_{n,m} H_{n-m} \bar{\phi}(t-nT) \phi(\tau-mT) \quad (4-138)$$

where the z-transform of $\{H_r\}$ is given by Eq. (4-61). Thus, the only difference between this optimum filter and that for the case where there is no random delay in the channel, is that this optimum filter uses the average pulse $\bar{\phi}$ to regenerate the signal rather than using the exact pulse ϕ . The weighting factors used as amplitudes for the regenerated signal are the same for both filters. Hence, linear estimators do not appear to be very useful for obtaining phase information.

(2) Random attenuation: We assume that the channel consists of a constant, but random, attenuation G_0 and additive white noise with PSD N_0 , and we consider the problem of estimating the transmitted signal x . For this problem, the correlation functions are

$$\begin{aligned} k_{yy}(t,s) &= \overline{G_0^2} k_{xx}(t,s) + N_0 \delta(t-s) \\ k_{xy}(t,s) &= \overline{G_0} k_{xx}(t,s), \end{aligned} \quad (4-139)$$

so that, from Theorem (4-7), we have the following solution for the optimum impulse-response function:

$$h_0(t, \tau) = \sum_{n,m} H_{n-m} \phi(t-nT) \phi(\tau-mT) \quad (4-140)$$

where the z-transform of $\{H_r\}$ is given by the formula

$$\begin{aligned} H(f) &= \overline{G_0} \Lambda(f) / (\overline{G_0^2} \Lambda(f) + N_0) \\ &= \overline{G_0} \Lambda(f) / (\overline{G_0}^2 \Lambda(f) + N(f)) \end{aligned} \quad (4-141)$$

where

$$N(f) \triangleq (\overline{G_0^2} - \overline{G_0}^2) \Lambda(f) + N_0 \quad (4-142)$$

From this second expression, we can interpret our solution as the solution to the problem where the channel attenuation is deterministic and equal to the mean attenuation \overline{G}_0 , and where the additive noise is the sum of the original white noise and the signal weighted by the variance of the channel attenuation $(\overline{G_0^2} - \overline{G_0}^2)$. This result parallels that of Maurer and Franks for optimum filtering of WSS signals through random channels [63].

Note that these results for PAM signals through random channels are valid for arbitrary CS processes if $A(f)$ is interpreted as the correlation matrix for the TSR of the CS signal.

vi) The video signal. An interesting example of a CS process that results from scanning is the video signal (Model (9) of Chapter II) whose Karhunen-Loeve TSR is discussed at length in Subsection 2d,2f of Chapter III. If we ignore the frame-to-frame correlation in the video signal, because of its relatively low-frequency nature, then the TSR correlation matrix is diagonal

$$\Lambda_{pq}(f) = \delta_{pq} \lambda_p \frac{1 - \alpha^2}{(1 - \alpha)^2 + 4\alpha \sin(\pi T f)} \quad (4-143)$$

where $\{\lambda_p\}$ are the eigenvalues, and where the basis function $\{\phi_p\}$ are the eigenfunctions (Eq. (3-27)):

$$\lambda_p = \frac{1/\pi f_0}{1 + \gamma_p^2},$$

$$\phi_p(t) = \begin{cases} \cos(\pi f_0 \gamma_p t), & p \text{ odd} \\ \sin(\pi f_0 \gamma_p t), & p \text{ even} \end{cases}, \quad t \in [-T/2, T/2]$$

where

$$\tan(\pi f_0 \gamma_p T) = \begin{cases} 1/\gamma_p, & p \text{ odd} \\ -\gamma_p, & p \text{ even} \end{cases} \quad (4-144)$$

In these expressions, α is the line-to-line correlation factor, and f_0 is the inverse-time-constant of the exponential correlation within a line.

Now, from Theorem (4-5), the optimum time-varying filter for the video signal in white noise has impulse-response function given by Eq. (4-56), (4-61b), and results in an average value for the periodically-varying minimum estimation-error given by Eq. (4-62b):

$$\begin{aligned} \langle e_0 \rangle &= \sum_{p=1}^{\infty} \int_{-1/2T}^{1/2T} \frac{N_0 \Lambda_{pp}(f)}{N_0 + \Lambda_{pp}(f)} df \\ &= \frac{N_0}{T} \sum_{p=1}^{\infty} \frac{\lambda_p}{[\lambda_p + N_0 \frac{1+\alpha}{1-\alpha}]^{1/2} [\lambda_p + N_0 \frac{1-\alpha}{1+\alpha}]^{1/2}}. \end{aligned} \quad (4-145)$$

As a specific example, we consider a square picture format with 500 lines and identical picture correlation along the horizontal and vertical:

$$\alpha = e^{-2\pi f_0 T/500} \quad (4-146)$$

where, for typical picture material, $0.7 \leq \alpha \leq 0.99$. We assume an effective single-sided bandwidth of $500/2T$ thereby obtaining a ratio of the time-averaged-signal-variance to the noise-variance equal

$$\rho = T/500N_0. \quad (4-147)$$

Now, since we are assuming that the receiver bandlimits the signal to the effective bandwidth of $500/2T$, then only the first 500 terms in the Karhunen-Loeve expansion will be significant [2], so that the infinite sum in Eq. (4-145) can be replaced with the first 500 terms. Using this truncated sum, we obtain (from digital-computer evaluation) the plots of minimum estimation-error vs. signal-to-noise ratio for various values of line-to-line correlation α shown in Fig. (4-17). The horizontal lines in this figure can be interpreted as the error which results from an ideal low-pass filter with cutoff frequency $500/2T$. Relative to this, the optimum time-varying filter provides a 10 dB improvement at low signal-to-noise ratio ($\rho=10$ dB) and high line-to-line correlation ($\alpha = .98$). As the signal-to-noise ratio increases and the correlation decreases, the improvement eventually becomes negligible. Furthermore, for all practical values of signal-to-noise ratio and line-to-line correlation, the improvement over the optimum time-invariant filter (not shown here) is negligible. This absence of significant improvement can be predicted in advance since the constraint of Eq. (4-146) on correlation renders the CS video signal "almost stationary". This low "degree of cyclostationarity" becomes obvious with the aid of Figures (1-3), (1-6) of Chapter I. The autocorrelation function for the video signal is the product of the functions shown in these two figures, and when Eq. (4-146) is satisfied with the order-of-magnitude-of- $\alpha = 1$, then the cyclic variations (along the lines parallel to the $t = s$ axis) of the blocks in the factor shown in Fig. (1-3) are almost totally attenuated by the narrow fences in the factor shown in Fig. (1-6). The cyclic character of this CS process will only be significant when the width of these fences becomes comparable

to (or greater than) the width of the blocks; i.e., when $f_0 T$ and α are both much less than unity, which signifies a video picture composed of nothing more than slightly correlated horizontal lines; i.e., a PAM-like signal with square pulses of one line-width. These results have been confirmed by digital-computer evaluation of $\langle \tilde{e}_0 \rangle$ and comparison with $\langle e_0 \rangle$, where $\langle \tilde{e}_0 \rangle$ is given by Eq. (4-60b), with $M = 500$.

vii) Time-division-multiplexed signals. We consider here the TDM₁ scheme (Mode! (3) of Chapter II) for time-division-multiplexing various (say M) signals $\{x_p\}$ together:

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^M x_p(t) w_p(t-nT) \quad (4-148)$$

where $\{w_p\}$ are the gate functions

$$w_p(t) \triangleq \begin{cases} 1 & \forall t \in [(p-1)T/M, pT/M] \triangleq T_p \\ 0 & \forall t \notin T_p \end{cases} \quad (4-149)$$

The Karhunen-Loeve TSR for this TDM₁ signal is discussed at length in Subsections 2d, 2f of Chapter III. If we assume that the component processes $\{x_p\}$ are uncorrelated and each have autocorrelation functions which satisfy the condition

$$k_p(t+rT-\tau) = \alpha_r^p k_p(t-\tau) \quad \forall t, \tau \in [0, T/M] \quad (4-150)$$

then, as shown in Subsection 2f, the TSR correlation matrix is diagonal, and if $\alpha_{n-m}^p = \delta_{nm}$ then, from a simple extension of Theorem (4-5) and Eq. (4-63), the optimum filter has impulse-response function

$$h_0(t, \tau) = \sum_{p=1}^M \sum_{q=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{\lambda_q^p}{N_0 + \lambda_q^p} \phi_q^p(t-nT-pT/M) \phi_q^p(\tau-nT-pT/M), \quad (4-151)$$

and results in minimum estimation-error

$$\langle e_o \rangle = \frac{1}{T} \sum_{p=1}^M \sum_{q=1}^{\infty} \frac{N_o \lambda_q^p}{N_o + \lambda_q^p}, \quad (4-152)$$

where $\{\phi_q^p; q = 1, 2, \dots\}$ and $\{\lambda_q^p; p = 1, 2, \dots\}$ are the eigenfunctions and eigenvalues for the kernel $k_p(\cdot, \cdot)$ on the interval $[0, T/M]$.

Now, if we choose triangular autocorrelation functions $\{k_p\}$ with widths $\{2\tau_p\}$ then, as shown in Subsection 2d, the eigenfunctions and eigenvalues are given by the formulas

$$\begin{aligned} \phi_q^p(t) &= \cos(\omega_q^p t) \\ \lambda_q^p &= 2k_p(0)/\tau_p (\omega_q^p)^2 \end{aligned} \quad (4-153)$$

where

$$\tan(\omega_q^p T/2M) = \frac{2M}{\omega_q^p T} - \frac{2M\tau_p}{T} [(2M/\omega_q^p T)^2 + 1]^{1/2} \quad (4-154)$$

provided that $T/2M \leq \tau_p \leq T(1-1/M)$.

Also, using a simple extension of Eq. (4-60b) results in the following formula for the minimum estimation-error resulting from the optimum time-invariant filter:

$$\langle \tilde{e}_o \rangle = \int_{-\infty}^{\infty} \frac{\frac{N_o}{T} \sum_{p=1}^M \sum_{q=1}^{\infty} |\phi_q^p(f)|^2 \lambda_q^p}{N_o + \frac{1}{T} \sum_{p=1}^M \sum_{q=1}^{\infty} |\phi_q^p(f)|^2 \lambda_q^p} df. \quad (4-155)$$

Evaluation of the performance index I for this TDM₁ signal requires numerical evaluation of Eqs. (4-152)-(4-155), and has not been carried out. However, judging from the TDM-of-PAM results of example ii), it

is expected that significant low-noise improvements will result if the individual variances and/or correlation-widths of the component processes $\{x_p\}$ differ significantly.

c) Examples employing the harmonic series representation.

i) Frequency-division-multiplexed signals. It was shown in the third example in the previous subsection that low-noise improvements of 3 dB can be obtained if optimum time-varying filters, rather than optimum time-invariant filters, are used for FDM signals. This result can be rederived using the HSR instead of the TSR previously used. However, we will employ the HSR here to study only the structure of the optimum time-varying filter.

As shown in Section 3d of Chapter III, the FDM signal

$$x(t) = \sum_{p=1}^M x_p(t) \cos(\omega_0 t + 2\pi p t/T) \quad (4-156)$$

$$\omega_0 = N2\pi/T$$

admits an HSR with correlation matrix

$$K_{pq}(f) = \begin{cases} \frac{1}{3} K_r(f), & |p| = |q| = r + N \\ 0, & \text{otherwise} \end{cases} \quad (4-157)$$

Thus, from Theorem (4-9), the impulse-response function for the optimum time-varying filter for the FDM signal in additive white noise is given by Eq. (4-65), where the matrix of transfer functions $\underline{H}(f)$ is given by Eq. (4-70), which reduces to the formula

$$H_{pq}(f) = \begin{cases} \frac{\frac{1}{2} K_r(f)}{K_r(f) + 2N_0}, & |p| = |q| = r + N \\ 0, & \text{otherwise} \end{cases} \quad (4-158)$$

as shown in Section 3e of Chapter III. This optimum filter has the general realization shown in Fig. (4-5). But due to the fact that the matrix \underline{H} displays the symmetry:

$$H_{pq} = \delta_{|p||q|} H_{pp} = \delta_{|p||q|} H_{p(-p)},$$

the structure of the filter can be reduced to that shown in Fig. (4-18), by combining paths. This structure has an obvious interpretation.

The bank of input modulators demultiplex the FDM signal into its component signals, which are then optimally filtered at baseband. The filtered signals are then remultiplexed together by the output bank of modulators. This is the same sequence of operations performed by the optimum time-varying filter for the TDM-PAM signal, and emphasizes the duality referred to in the last subsection.

ii) Amplitude modulated signals. Amplitude modulation is one of the most frequently used signal formats for communication. In fact, the FDM signal considered in the previous example is nothing more than the sum of M AM signals. Hence, the results for FDM with $M = 1$ are valid for AM signals. That is, for the AM signal (Model (1) of Chapter II)

$$x(t) = x_0(t) \cos(2\pi t/T), \quad (4-159)$$

where the PSD of x_0 is bandlimited to $[-1/2T, 1/2T]$, the optimum time-varying filter has the structure of a single path of the M -path filter for FDM shown in Fig. (4-18), and provides a 3 dB improvement in performance over that of the optimum time-invariant filter. Paralleling the operations performed by the optimum filters for FDM and TDM-PAM, this optimum time-varying filter demodulates the AM signal, optimally

filters it at baseband, and then remodulates the filtered signal.

iii) Sinusoid with random amplitude and phase. As an example of the general results obtained in Section 5c for LMS smoothing of random periodic signals in noise, we consider here the random sinusoid

$$x(t) = a \sin(2\pi t/T) + b \cos(2\pi t/T), \quad (4-160)$$

where a , b are zero-mean random variables with variances σ_a^2 , σ_b^2 and crosscovariance σ_{ab}^2 . This process admits a first order HSR with correlation matrix elements

$$K_{-1-1} = K_{11}^* = \sigma_a^2/4 - \sigma_b^2/4 + j\sigma_{ab}^2/2$$

$$K_{-11} = K_{1-1} = \sigma_a^2/4 + \sigma_b^2/4$$

$$K_{0\pm 1} = K_{\pm 10} = K_{00} = 0. \quad (4-161)$$

Now, from Theorem (4-10) the optimum time-varying smoother for this signal in additive white noise has impulse-response function given by Eq. (4-74), and results in the time-averaged minimum estimation-error given by Eq. (4-76), which reduces to

$$\langle e_o \rangle = (2N_o/NT) \frac{A + B}{\Lambda + 2B + 4N_o/NT},$$

where

$$\Lambda = \frac{NT}{N_o} (\sigma_a^2 \sigma_b^2 - \sigma_{ab}^4) \quad (4-162)$$

$$B = \sigma_a^2 + \sigma_b^2.$$

From Theorem (4-11), the optimum time-invariant smoother has impulse response function given by Eq. (4-77), and results in the estimation error given by Eq. (4-78), which reduces to

$$\langle \tilde{e}_0 \rangle = (2N_0/NT) \frac{B}{B + 4N_0/NT} \quad (4-163)$$

From these last two formulas, it is clear that for low-noise conditions, the performance index becomes

$$I \approx \frac{A + 2B}{A + B}$$

and is a maximum of 3 dB when $A = 0$. But $A = 0$ iff $\sigma_a \sigma_b = \sigma_{ab}^2$ iff a and b are proportional iff the phase of the random sinusoid is deterministic. This result is exactly what one would expect: The time-varying filter outperforms the time-invariant filter by the greatest margin when the sinusoid displays its highest degree of cyclostationarity; i.e., when its phase is completely known.

d) Examples employing the Fourier series representation.

i) Amplitude-modulated signals. In the previous two subsections we employed the TSR and HSR to derive the structure of the optimum time-varying filter for AM signals, and to show that the performance index is a maximum of 3 dB for low-noise conditions. These results can also be derived using the FSR as briefly outlined here: The autocorrelation function for the AM signal

$$x(t) = x_0(t) \cos(2\pi t/T)$$

where x_0 is bandlimited to $[-1/4T, 1/4T]$ admits a second order FSR with coefficient functions

$$\begin{aligned}
C_0(f) &= \frac{1}{4} K_0(f-1/T) + \frac{1}{4} K_0(f+1/T) \\
C_1(f) &= C_{-1}(f) = 0 \\
C_2(f) &= C_{-2}(f) = \frac{1}{4} K_0(f)
\end{aligned} \tag{4-164}$$

where K_0 is the PSD for x_0 .

Now, from Theorem (4-13), the optimum time-varying filter admits the second order FSR of Eqs. (4-83)-(4-85). Substituting Eq. (4-164) into these formulas yields formulas for the transfer functions for the time-invariant filters in the optimum time-varying filter structure of Fig. (3-14a). With these formulas, this structure can be reduced to the simple demodulator-filter-modulator structure identified in example ii) of the last subsection.

Also, this optimum time-varying filter results in the minimum estimation-error given by Eq. (4-88) of Theorem (4-13). Substitution of Eq. (4-164) into this formula yields the following expected formula:

$$\langle e_0 \rangle = \int_{-1/2T}^{1/2T} \frac{N_0 K_0(f)}{2N_0 + K_0(f)} df. \tag{4-165}$$

Note that the above result is valid as long as x_0 is bandlimited to $[-1/2T, 1/2T]$, but the solution method of Theorem (4-13) is only valid for x_0 bandlimited to $[-1/4T, 1/4T]$.

ii) Bandpass PAM. Consider the following bandpass PAM signal

$$x(t) = \sqrt{2T} \sum_n a_n \phi(t-nT),$$

where

$$\phi(t) = \sqrt{1/2T} \cos(\omega_0 t) \frac{\sin(\pi t/4T)}{(\pi t/4T)}, \quad (4-166)$$

where ω_0 is an integer-multiple of $2\pi/T$. Now, if we consider the $\{a_n\}$ to be the sample values $\{x_0(nT)\}$ of a process which is bandlimited to $[-1/4T, 1/4T]$, then Eq. (4-166) can be reexpressed, using the sampling theorem (Theorem (2-2)), as

$$x(t) = 2x_0(t)\cos(\omega_0 t)$$

and is therefore equivalent to the AM signal of the previous example.

Hence, the optimum filter and estimation-error derived there are valid here. However, the most meaningful filter structure results from using the TSR (this structure consists of one path of the M-path structure of Figure (4-16)) rather than the FSR (or the HSR). In fact, the solution obtained using the TSR is valid even if the a_n are not the sample values of a process which is bandlimited.

iii) Improvement for general PAM. In subsection 6b, we presented the general formulas of Eq. (4-132) for estimation-errors for arbitrary PAM signals in additive white noise. We will now reexpress these formulas in terms of FSR coefficient functions in order to gain insight into the relationship between improved performance and the FSR coefficient functions.

Substituting the TSR of Eq. (3-2) for the autocorrelation function for the PAM process into the formula of Eq. (3-101) for the p^{th} FSR coefficient function yields the formula

$$c_p(t) = \frac{1}{T} \sum_r A_r \theta(t-rT) e^{-j2\pi p t/T}$$

where

$$\theta(t) \triangleq \int_{-\infty}^{\infty} \phi(t+\tau)\phi(t)dt.$$

Thus, the Fourier transform C_p of c_p takes the form

$$C_p(f) = \frac{1}{T} \Lambda(f) \phi(f+p/2T) \phi^*(f-p/2T). \quad (4-167)$$

Now, from this formula, we obtain the relation

$$\frac{1}{T} \Lambda(f) |\phi(f+p/T)|^2 = C_0(f+p/T) = |C_p(f+p/T)|^2 / C_0(f). \quad (4-168)$$

Substituting this into the formulas of Eq. (4-132) yields the new formulas

$$\begin{aligned} \langle c_o \rangle &= \sum_p \int_{-1/2T}^{1/2T} \frac{N_o C_0(f+p/T)}{N_o + J_p(f) C_0(f+p/T)} df \\ \langle \tilde{c}_o \rangle &= \sum_p \int_{-1/2T}^{1/2T} \frac{N_o C_0(f+p/T)}{N_o + C_0(f+p/T)} df \\ J_p(f) &\triangleq \sum_q \frac{|C_q(f+q/T)|^2}{C_0(f) C_0(f+p/T)}. \end{aligned} \quad (4-169)$$

Now, from these formulas, it is clear that for every value of the index p for which $J_p(f) > 1$, $\forall f \in [-1/2T, 1/2T]$, there is a contribution to the overall improvement-in-performance (of the optimum time-varying filter over the optimum time-invariant filter). Furthermore, $J_p(f)$ is a measure of the relative magnitudes of the zeroth and p th order FSR coefficient functions. Thus, we have here a positive relationship between large higher order FSR coefficient functions, and large improvements.

e) Examples employing the inverse-operator-feedback-system method.

In subsec. 5e we presented four classes of optimum filtering problems to which the inverse-operator-feedback-system (IOFS) method of solution

applies. We now illustrate these with solutions for specific problems.

i) Periodically attenuated WSS signal in WSS noise. Consider the problem of optimally filtering a bandpass WSS signal with PSD

$$K_{xx}(f) = \begin{cases} \sigma_x^2 T/2, & ||f| - f_0| \leq 1/2T \\ 0, & \text{otherwise} \end{cases} \quad (4-170)$$

which has been transmitted through a channel with slowly-varying periodic attenuation, $1 + a \cos(2\pi t/T)$, and additive WSS white noise with PSD N_0 .

This problem belongs to class (1) of Subsec. 5e with $M = 2Tf_0 \gg 1$

(assumed to be an integer), and $N = 1$ in Eq. (4-97). Thus, the impulse-response function for the optimum time-varying filter is, from Eqs. (4-98), (4-99), (4-101),

$$h_0(t, \tau) = \frac{1 + a \cos(2\pi \tau/T)}{N_0 + \sigma_x^2 T(1+a^2/2)/2} k_{xx}(t-\tau) \quad (4-171)$$

and has the realization shown in Fig. (4-19). Furthermore, the time-averaged value of the minimum estimation-error is, from Eq. (4-29)

$$\begin{aligned} \langle e_0^2 \rangle &= \sigma_x^2 - \frac{1}{N_0 + \sigma_x^2 T(1+a^2/2)/2} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{\infty} (1 + a \cos(2\pi s/T))^2 k_{xx}^2(t-s) ds dt \\ &= \frac{\sigma_x^2}{1 + \rho(1+a^2/2)} \end{aligned} \quad (4-172)$$

where ρ is the signal-to-noise ratio ($\rho = \sigma_x^2 T / 2N_0$).

The estimation-error which results from the optimum time-invariant filter (i.e., that for the problem where the periodic attenuation is replaced with its average value) is, from Eq. (4-27)

$$\langle \tilde{e}_0^2 \rangle = \frac{\sigma_x^2}{1 + \rho} \quad (4-173)$$

Hence, there is a low-noise improvement factor of $I = (1+a^2/2)$, and $I < 3/2$, assuming $|a| < 1$.

It is interesting to note that the optimum filter, Fig. (4-19) decomposes into the cascade connection of a periodic multiplier, which attempts to equalize the periodic attenuation, and a Wiener filter which optimally filters the equalized signal. It should be emphasized, however, that the multiplier is not an equalizer in the true sense, and the use of an exact equalizer would result in inferior performance.

ii) AM-type CS signal in WSS noise. For the class of optimum-filtering problems where a CS signal, which is the product of a deterministic T-periodic signal q and a WSS process z with PSD

$$K_{zz}(f) = \begin{cases} \sigma_z^2 T, & |f| \leq 1/2T \\ 0, & \text{otherwise,} \end{cases} \quad (4-174)$$

is transmitted through a channel with additive WSS white noise with PSD N_0 (class (2) of Subsection 5e), the optimum time-varying filter is easily shown (using Eqs. (4-98), (4-99), (4-103)) to have impulse-response function

$$h_0(t, \tau) = q(t)q(\tau)k_{xx}(t - \tau) / (Q\sigma_x^2 T + N_0). \quad (4-175)$$

Thus, the minimum estimation-error is, from Eq. (4-32),

$$\langle e_0 \rangle = \frac{Q\sigma_z^2 N_0 / T}{Q\sigma_z^2 + N_0 / T},$$

where

$$Q \triangleq \sum_n |q_n|^2, \quad (4-176)$$

where $\{q_n\}$ are the Fourier coefficients for the T-periodic function q .

Also, from Eq. (4-27), the optimum time-invariant filter results in the estimation-error

$$\langle \tilde{e}_0 \rangle = \sum_n \frac{\sigma_z^2 |q_n|^2 N_0 / T}{\sigma_z^2 |q_n|^2 + N_0 / T}$$

(where we have used

$$K_{xx}(f) = \sum_n |q_n|^2 K_{xx}(f+n/T). \quad (4-177)$$

Now, for low-noise conditions, we have

$$\langle e_0 \rangle = N_0 / T.$$

$$\langle \tilde{e}_0 \rangle > N_0 2N / T \quad (4-178)$$

where $N_0 \ll \sigma_z^2 |q_n|^2 T$ for $|n| \leq N$, so that there is a low-noise improvement factor of $1 > 2N$. Hence, there can be significant improvement if the periodic factor q has high harmonic content as, for example, in a rectangular gating function.

One should be careful in interpreting this result. For example, if q is the rectangular gating function shown in Fig. (4-20a), then one might reason that the optimum time-varying filter obviously performs much better than the optimum time-invariant filter since the former passes zero noise to the output during the majority of every T -second interval (when the signal x is itself zero), whereas the latter can not behave in such a time-varying fashion. That this reasoning is invalid is clear when we consider the new gating function shown in Fig. (4-20b). The improvement factor is the same for this gating function as it is for that in Fig. (4-20a), yet the above argument obviously does not apply, since the optimum time-

varying filter now passes zero noise to the output only during a very small portion of every T -second interval. Hence, it must be concluded that the improvement is simply due to the fact that the time-invariant filter can not reproduce the sharp features in the gated signal without passing high-frequency noise, whereas the time-varying filter easily reproduces the sharp features by employing a gate of its own.

iii) WSS signal in white CS noise. Consider the class of optimum-filtering problems where a WSS signal with PSD

$$K_{xx}(f) = \begin{cases} \sigma_x^2 T, & |f| \leq 1/2T \\ 0, & \text{otherwise} \end{cases} \quad (4-179)$$

is transmitted through a channel with additive white CS noise with autocorrelation function

$$k_{nn}(t,s) = q^2(t)\delta(t-s), \quad (4-180)$$

These problems belong to class (3) of Subsection 5e. Thus, the impulse-response function for the optimum time-varying filter is given by Eqs. (4-98), (4-99), (4-105):

$$h_o(t,\tau) = q^{-2}(\tau) \frac{1}{1 + p\sigma_x^2 T} k_{xx}(t-\tau), \quad (4-181)$$

and the resultant estimation-error is, from Eq. (4-29),

$$\begin{aligned} \langle e_o \rangle &= \sigma_x^2 - \frac{1}{T} \int_{-T/2-\infty}^{T/2} \int_{-\infty}^{\infty} q^{-2}(s) k_{xx}^2(t-s) ds dt \\ &= \frac{\sigma_x^2}{1 + p\sigma_x^2 T}, \end{aligned} \quad (4-182)$$

where

$$P \triangleq \frac{1}{T} \int_{-T/2}^{T/2} q^{-2}(t) dt.$$

The estimation-error which results from the optimum time-invariant filter is, from Eq. (4-25),

$$\langle \tilde{e}_0^2 \rangle = \frac{\sigma_x^2}{1 + \sigma_x^2 T/Q} \quad (4-183)$$

where

$$Q \triangleq \frac{1}{T} \int_{-T/2}^{T/2} q^2(t) dt.$$

Hence, the low-noise improvement factor is $I = QP$. For example, if the noise intensity $q^2(t)$ switches back and forth every $T/2$ seconds from N_0 to aN_0 ($a \leq 1$), then

$$QP = \frac{1}{4} (1 + a)(1 + 1/a), \quad (4-184)$$

and the improvement exceeds 3 dB for $a < 1/5$.

iv) Doppler-shifted Poisson pulse-train. We consider here the problem of optimally filtering a Doppler-shifted Poisson pulse-train in additive white noise. We model the signal x as an asynchronous PAM signal which has been subjected to a periodic time-scale transformation $t' = t + q(t)$ (Model (8) of Chapter II):

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \phi(t+q(t)-t_n), \quad (4-185)$$

where the original (before Doppler) pulse-occurrence-times $\{t_n\}$ form an order sequence distributed according to the Poisson counting process with constant rate parameter μ_0 , and where $\{a_n\}$ is a zero-mean WSS sequence.

The PSD for the original process without Doppler ($q(t) \equiv 0$) is

$$K_{zz}(f) = \mu_0 \sigma_a^2 |\phi(f)|^2, \quad (4-186)$$

so that if the pulse ϕ is a sinc pulse with Fourier transform

$$\phi(f) = \begin{cases} T, & |f| \leq 1/2T \\ 0, & \text{otherwise,} \end{cases} \quad (4-187)$$

then this optimum-filtering problem belongs to class (4) of Subsec. 5e where $g(t) = t + q(t)$. Thus, the impulse-response function for the optimum time-varying filter is, from Eq. (4-98), (4-99), (4-107):

$$h_o(t, \tau) = \frac{1}{P \mu_0 \sigma_a^2 T^2 + N_0} k_{zz}(g(t) - g(\tau)), \quad (4-188)$$

where

$$P \triangleq \frac{1}{T} \int_{-1/2T}^{1/2T} 1/g'(t) dt$$

and has the realization shown in Fig. (4-21). The structure of this optimum filter has the interesting configuration consisting of an inverse time-scale transformation, followed by a time-invariant Wiener filter and a time-scale transformation at the output. Thus, the filter removes the periodic fluctuations from the CS signal, optimally filters the resultant WSS signal, and then re-inserts the periodic fluctuations. This sequence of signal processing operations parallels the demodulator-filter-modulator, and demultiplexor-filter-multiplexor sequences derived in many of the previous examples.

The time averaged value of the minimum estimation-error which results from the optimum time-varying filter is, from Eq. (4-32),

$$\begin{aligned}
 \langle e_o \rangle &= \frac{N_o}{P\mu_o\sigma_a^2T^2 + N_o} \frac{1}{T} \int_{-1/2T}^{1/2T} k_{zz}(g(t)-g(t))dt \\
 &= \frac{\mu_o\sigma_a^2T^2N_o/P}{\mu_o\sigma_a^2T^2 + N_o/P} \quad (4-189)
 \end{aligned}$$

v) Time-division-multiplexed signals. In Chapter II we introduced the information lossless TDM₂ scheme (Model (5)) as an alternative to the information lossy TDM₁ scheme (Model (3)) for time-division-multiplexing various signals together. The TDM₂ signal is formed by periodically time-compressing the component processes into non-overlapping time-slots, so that when the components are added they will not interfere with each other (See Fig. (2-7) of Chapter II). If the M component processes are statistically independent, then it can be shown that the optimum time-varying filter for the TDM₂ signal in additive white noise is a parallel connection of M independent filters, the pth of which is the optimum time-varying filter for the pth time-compressed component process. These M filters include periodic gates at their inputs so that the pth filter admits only the pth component signal (in white noise).

Now, the impulse-response function for the pth optimum filter must satisfy the orthogonality condition

$$\int_{-\infty}^{\infty} k_{x_p x_p}(\tau, \sigma) h_p(t, \sigma) d\sigma + N_o h_p(t, \tau) = k_{x_p x_p}(t, \tau)$$

where

$$k_{x_p x_p}(t, s) = k_{z_p z_p}(t + q(t + (M-p)T/M) - s - q(s + (M-p)T/M) \cdot$$

$$w(t - (p-1)T/M)w(s - (p-1)T/M)$$

where z_p is the original component process from which the time-compressed version x_p is obtained, and where w is the periodic gate function

$$w(t) \triangleq \begin{cases} 1, & 0 \leq |t-nT| \leq T/M, \quad \forall n \\ 0, & \text{otherwise,} \end{cases}$$

and where q is the periodic time-scale compression function shown (as p) in Fig. (2-6) of Chapter II.

This optimum-filtering problem is a combination of the types of problems in classes (2) and (4) of Subsection 5e, and can therefore be solved in some cases using the IOFS method of that subsection.

f) Summary and Conclusions. In concluding this chapter on optimum filtering of cyclostationary processes, we emphasize the intuitively pleasing result that optimum time-varying filters for cyclostationary processes can always be decomposed into a sequence of three signal processing operations; 1. The cyclostationary process is decomposed into one or more (jointly) stationary processes. 2. The stationary component-processes are passed through time-invariant filters (often Wiener filters). 3. The filtered components are recombined to form the optimally filtered cyclostationary process. In most of the examples presented, this sequence of signal processing operations takes one of the following specific forms: demodulator-filter-modulator, demultiplexor-filter-multiplexor, inverse time-scale transformation--filter--time-scale transformation. For those particular cyclostationary processes which decompose into a single stationary process, the three operations are more accurately described as: 1. Removal of periodic fluctuations. 2. Time-invariant filtering. 3. Re-insertion of periodic fluctuations.

In fact, the optimum time-varying filters derived for FSK, PAM, TDM-PAM, AM, and FDM signals in additive white noise all consist of demultiplexor-demodulator--time-invariant Wiener filter--modulator-multiplexor structures.

Now, if the optimum filter is to be employed in a receiver, then it will be followed by a demultiplexor-demodulator. But the same "baseband" signal can be obtained more directly by simply eliminating the third operation of modulation and multiplexing from the optimum time-varying filter, thereby significantly reducing the complexity of the overall receiver.

Hence, the optimum linear receiver is a demultiplexor-demodulator followed by a time-invariant filter(s). It is interesting to note that this optimum linear receiver employs the same two types of operations as the suboptimum receiver which ignores the cyclic character of the cyclostationary signals, but in reverse order; i.e., the suboptimum receiver consists of an optimum time-invariant filter followed by a demultiplexor-demodulator.

We state this very practical result here as a theorem to be referred to as the Optimum Linear Receiver Theorem:

THEOREM(4-14): Consider the class of signals which are obtained by linear modulation and multiplexing of WSS waveforms and data sequences.

If the additive noise to which these signals are subjected during transmission is equivalent²⁹ to WSS white noise after demultiplexing and demodulating, then the optimum (LMS) linear receiver consists of the demultiplexor-demodulator followed by time-invariant (continuous- or discrete-time) Wiener filters.

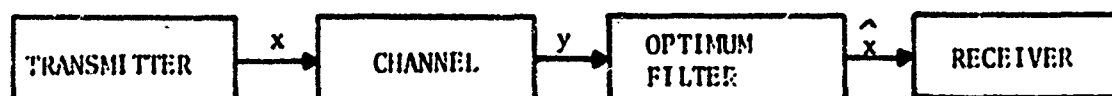
²⁹ In this theorem, two noises are "equivalent" if they result in the same orthogonality condition for LMS estimation.

The proof of this theorem is straightforward provided that the linear modulation-multiplexing transformation is invertible--a requirement which must be satisfied in practice.

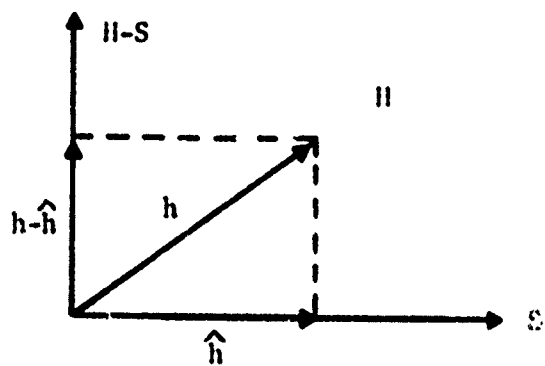
The essence of this theorem is, as discussed above, the fact that the time-invariant filtering must follow--not precede--the demultiplexing and demodulation.

Another result worth emphasizing is that cyclostationary processes whose mean and variance are constant can still possess a high degree of cyclostationarity in the sense that optimum time-varying filtering can yield substantially lower mean-squared error than optimum time-invariant filtering.

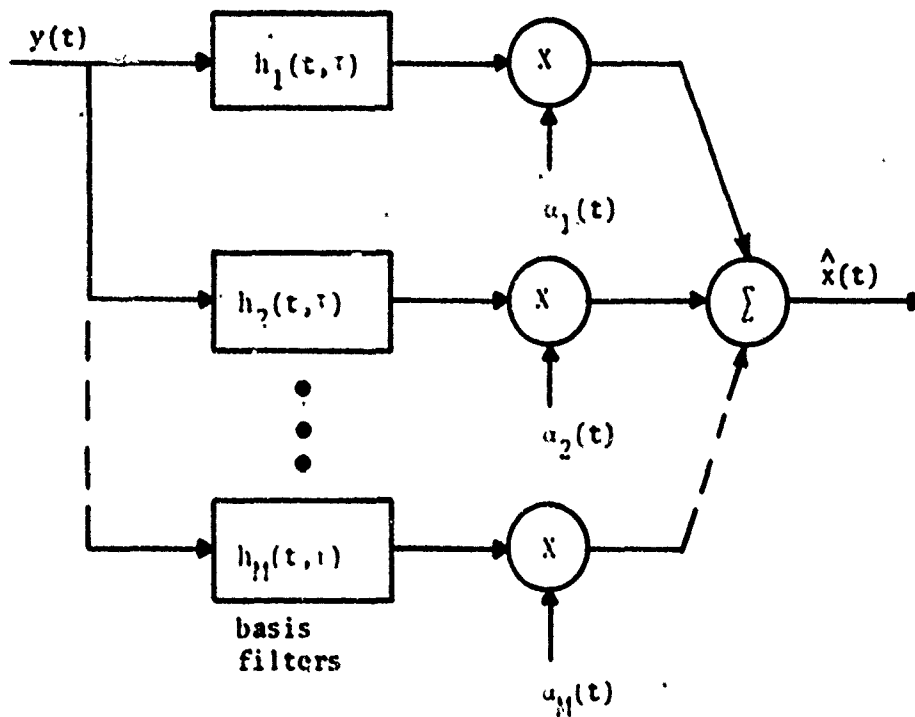
A third interesting result is that optimum time-varying filters for cyclostationary processes appear to yield the greatest reduction in mean-squared error (below that resulting from optimum time-invariant filters) when the signal-to-noise ratio is high, but can yield substantial reductions even at signal-to-noise ratios as low as unity.



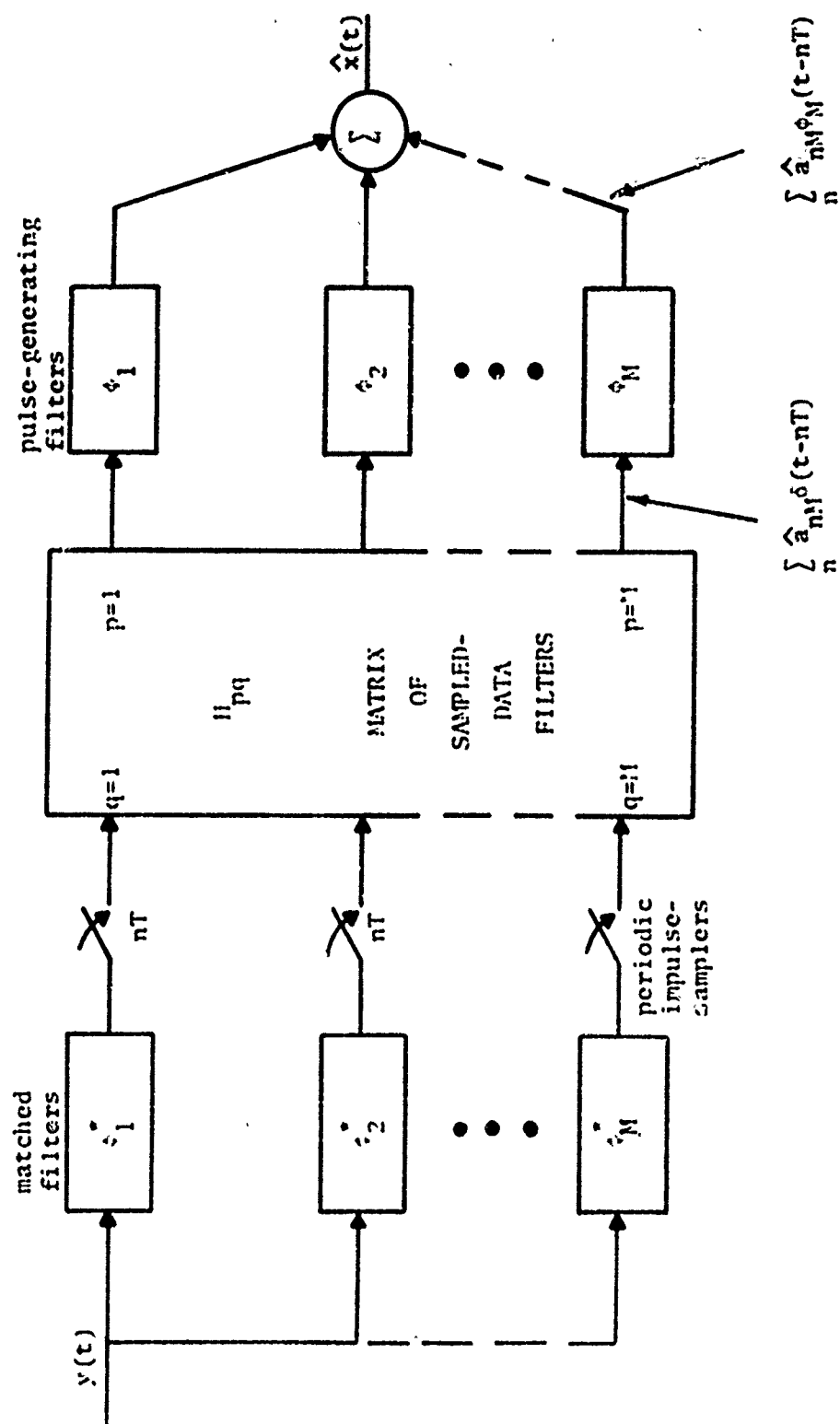
Figure(4-1) Communication-system diagram.



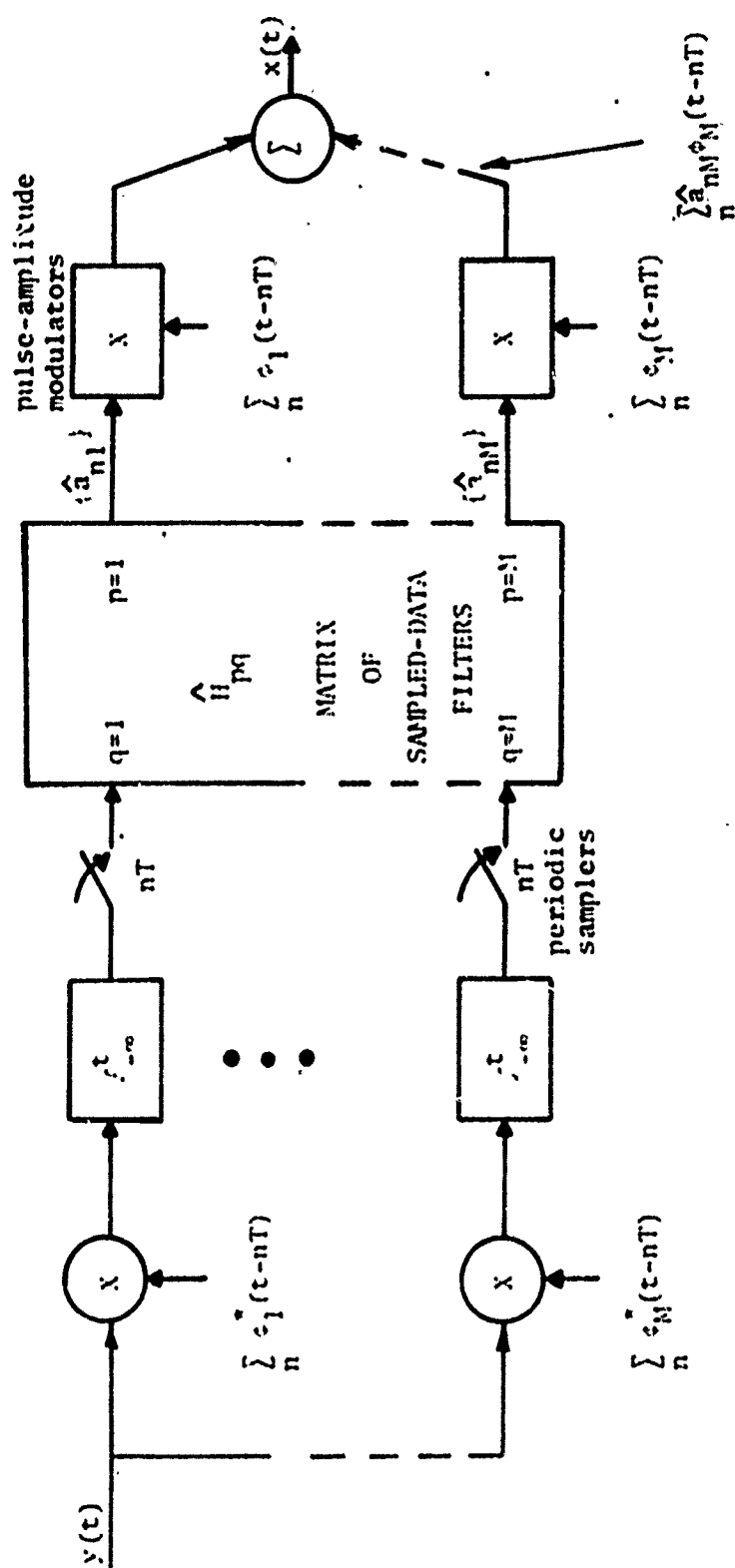
Figure(4-2) Orthogonal projection in two-dimensional Euclidean space.



Figure(4-3) Structure of an M-dimensional optimum filter
 where $\{h_i\}$ are chosen a-priori, and $\{a_i\}$ are optimized.



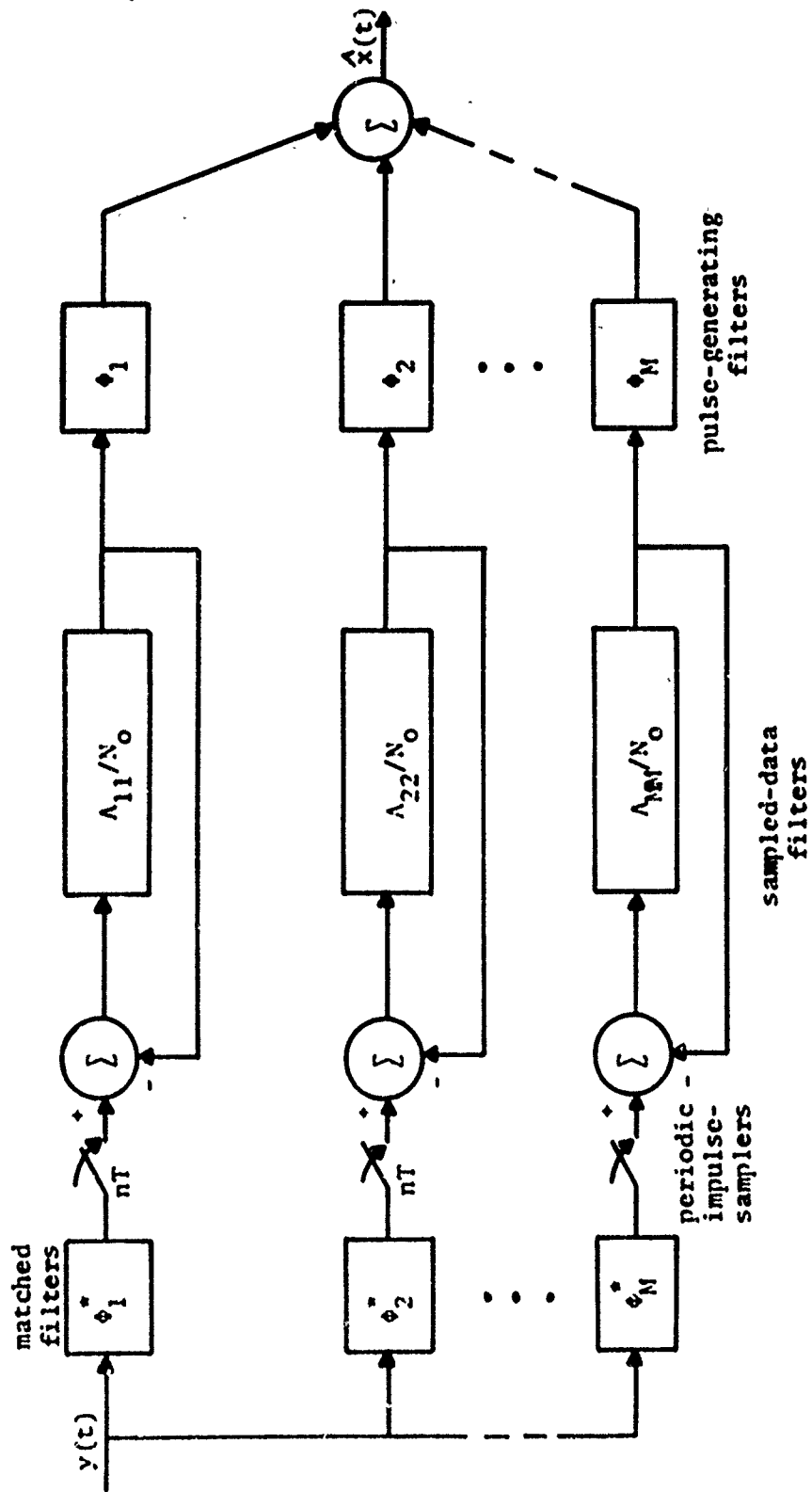
Figure(4-4a) Structure of optimum filter derived using M^{th} order TSR's, as in Theorems(4-5,6,7). Note, for the filter of Theorems(4-6,7), the transfer functions for the input filters are $\{\theta_p^*\}$.



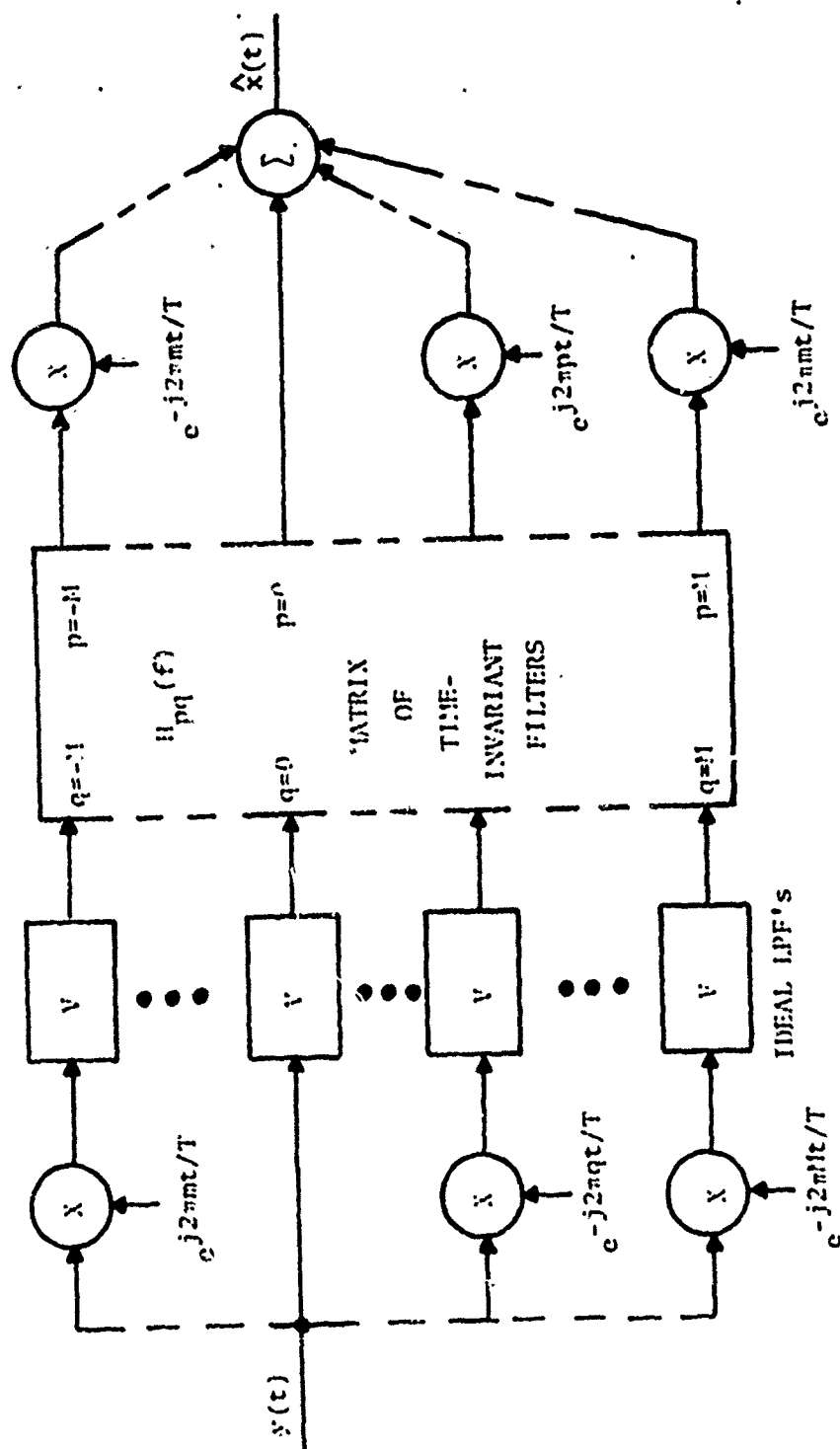
Figure(4-4b) Alternate structure for realization of optimum filters derived using N^{th} order.

TSR's. Note, for Theorems(4-6,7), the input modulating functions are composed of $\{a_i^*\}$.

$$(\hat{H}_{pq}(\epsilon) = H_{pq}(\epsilon)/(1 + e^{-j2\pi\epsilon T})$$

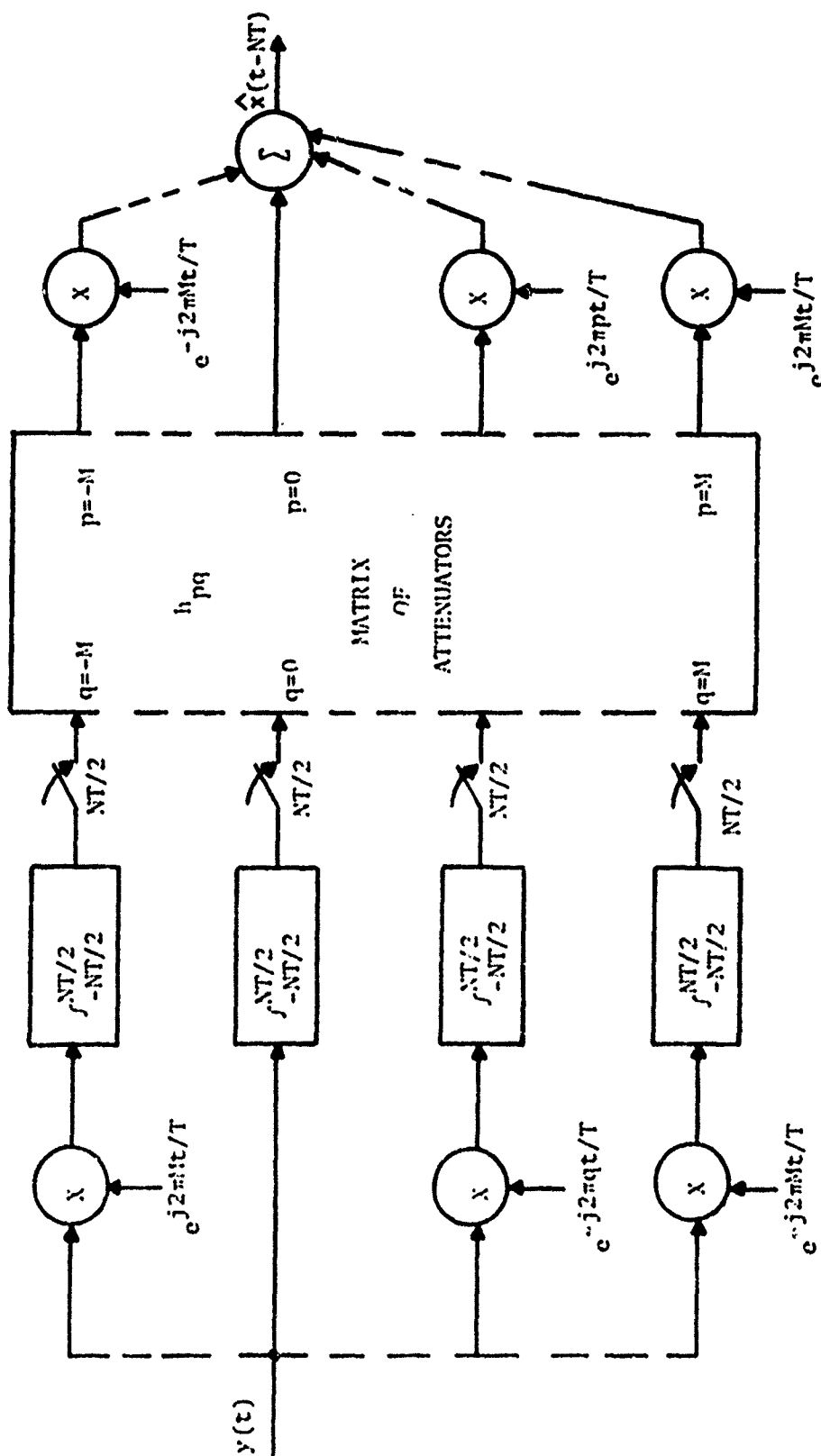


Figure(4-4c) Structure of optimum filter for T-CS signal in additive white noise when the signal's M^{th} order TSR correlation matrix $\underline{\Lambda}(f)$ is diagonal (Sec Eq. (4-61).)



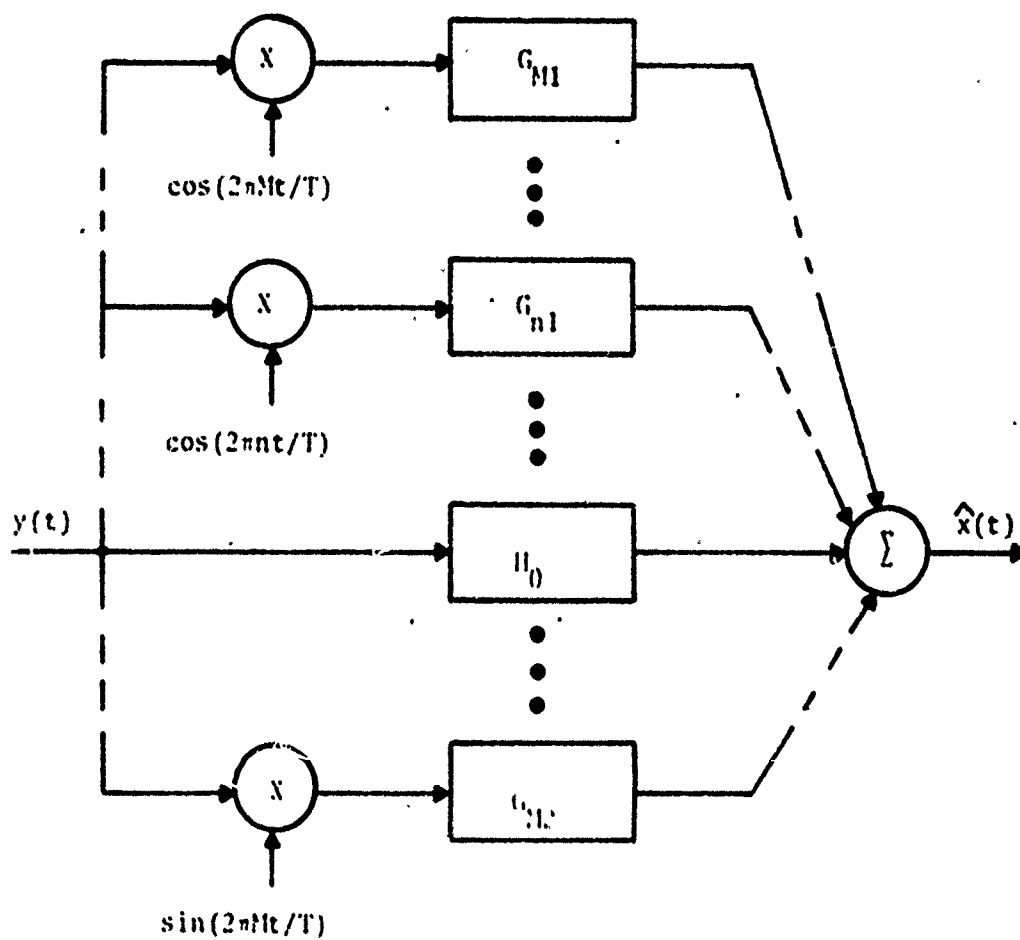
Figure(4-5) Structure of optimum filter derived using an N^{th} order IIR. (See Theorem(4-9).)

V is an ideal low-pass filter with cutoff frequency $1/2T$.



Figure(4-6) Structure of optimum fixed-interval smoother for random periodic signals in noise.

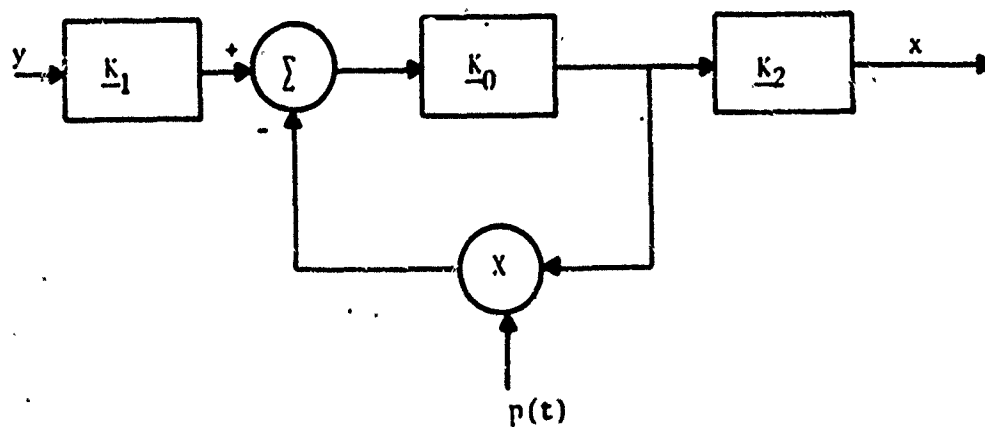
(See Theorem(4-10).)



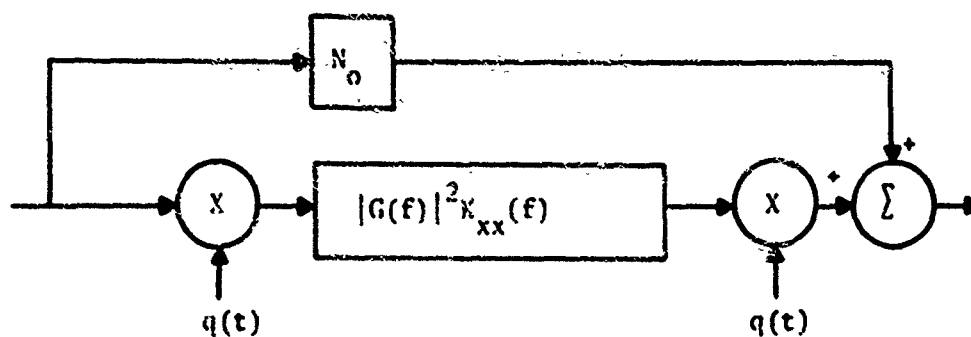
Figure(4-7) Structure of optimum filter derived using an N^{th} order FSR. (See Theorem(4-13).)



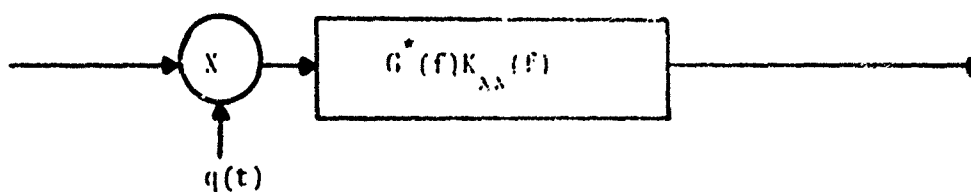
Figure(4-8) System diagram of optimum filter. (\underline{K}_{yy}^{-1} , \underline{K}_{xy} are T-periodic linear systems.)



Figure(4-9) Feedback configuration for optimum filter. (\underline{K}_0 is a time-invariant linear system, $p(\cdot)$ is a T-periodic function, and $\underline{K}_1, \underline{K}_2$ are T-periodic linear systems.)

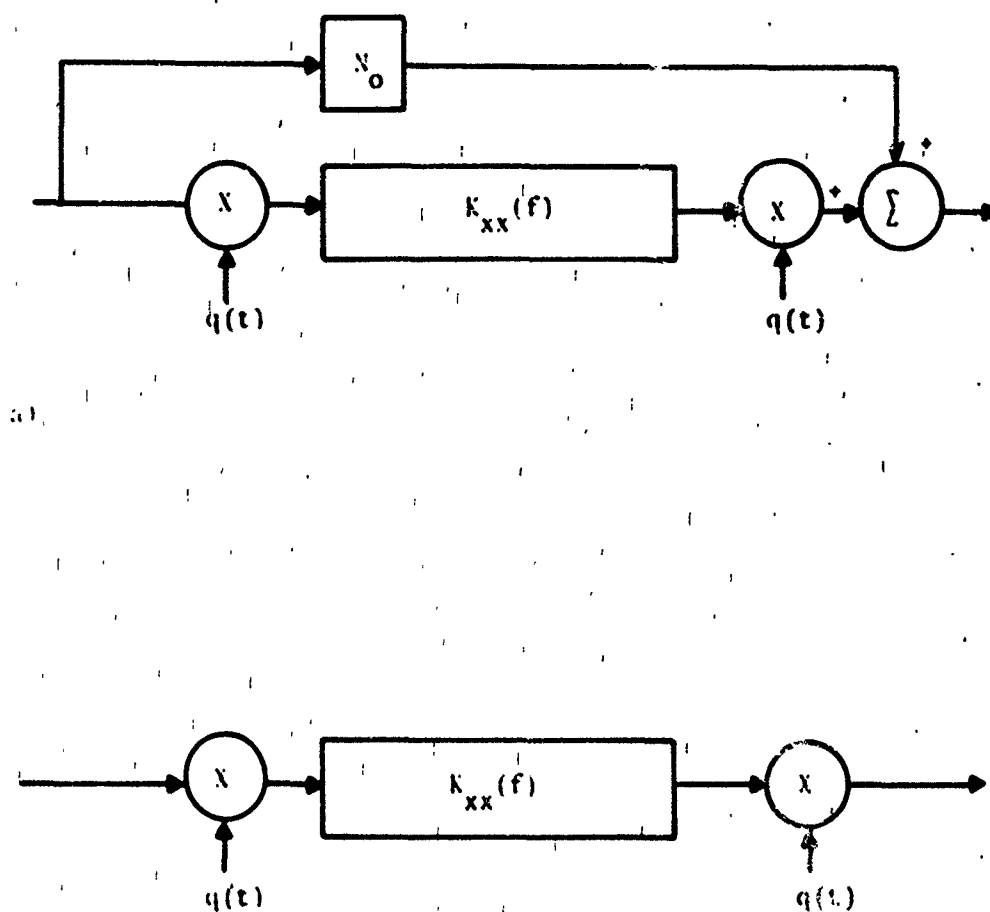


a)

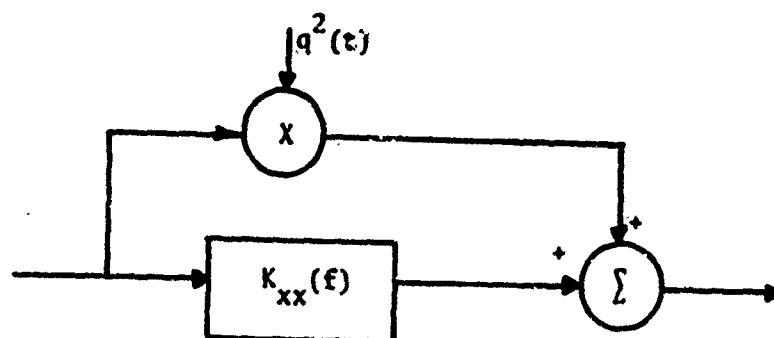


b)

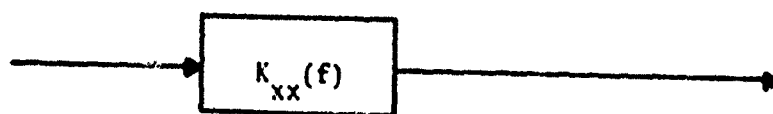
Figure(4-10) a) Realization of the operator K_{xy} . b) Realization of the operator K_{xy} . (For example 1.)



Figure(4-11) a) Realization of the operator K_{yy} . b) Realization of the operator K_{xy} . (For example 2.)

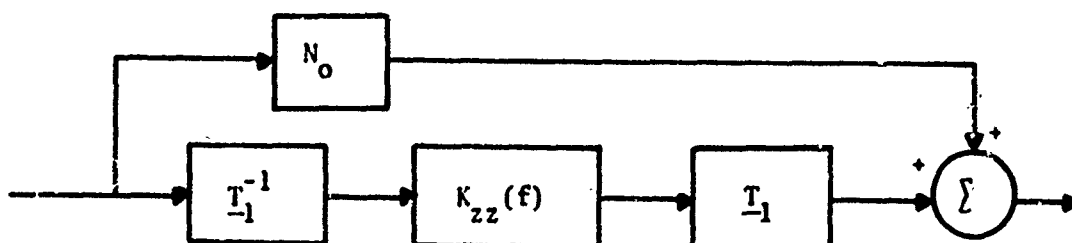


a)

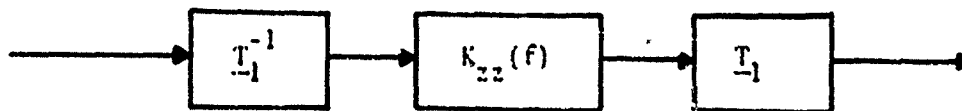


b)

Figure(4-12) a) Realization of the operator \underline{K}_{yy} . b) Realization of the operator \underline{K}_{xy} . (For example 3.)

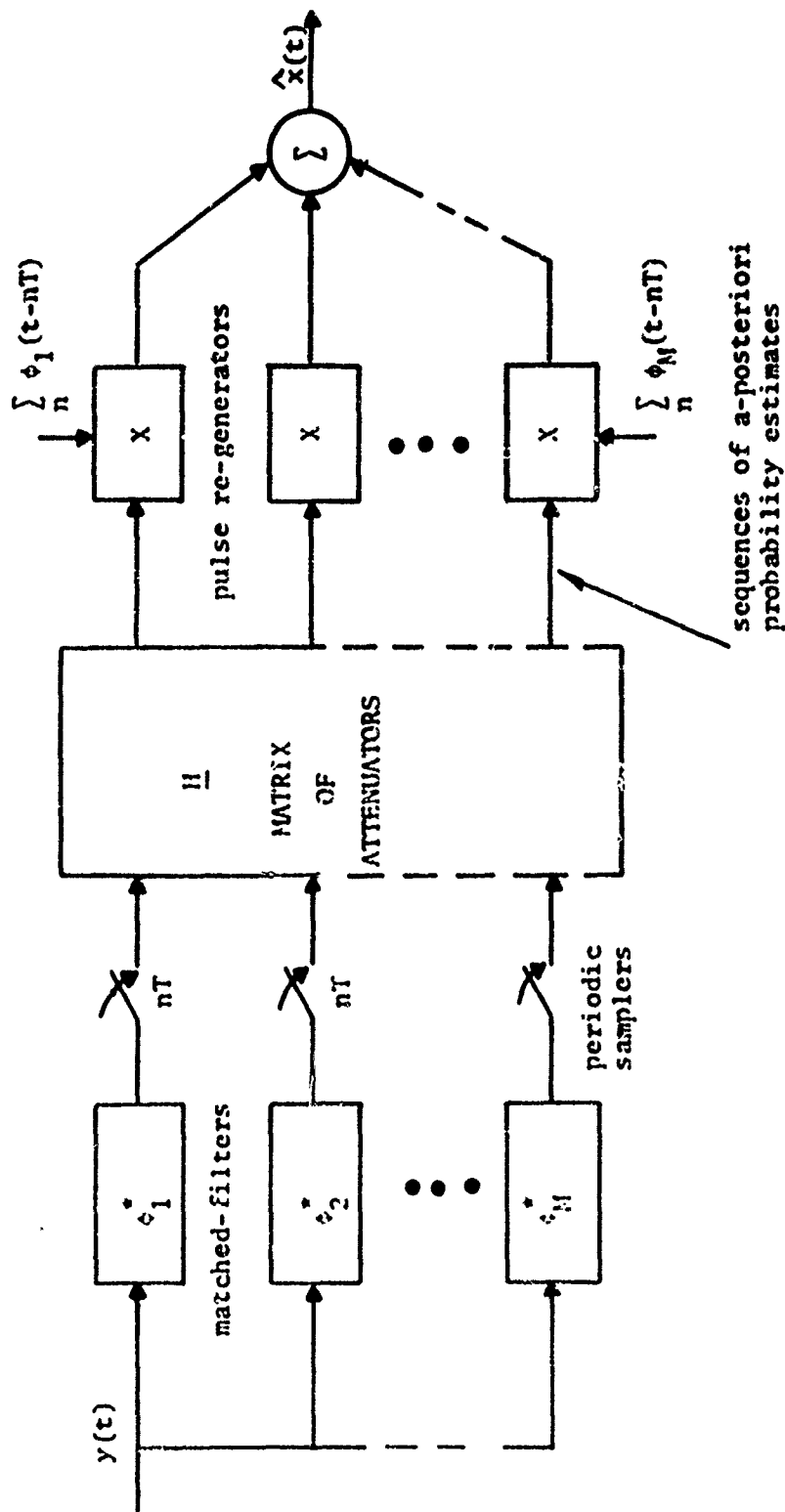


a)



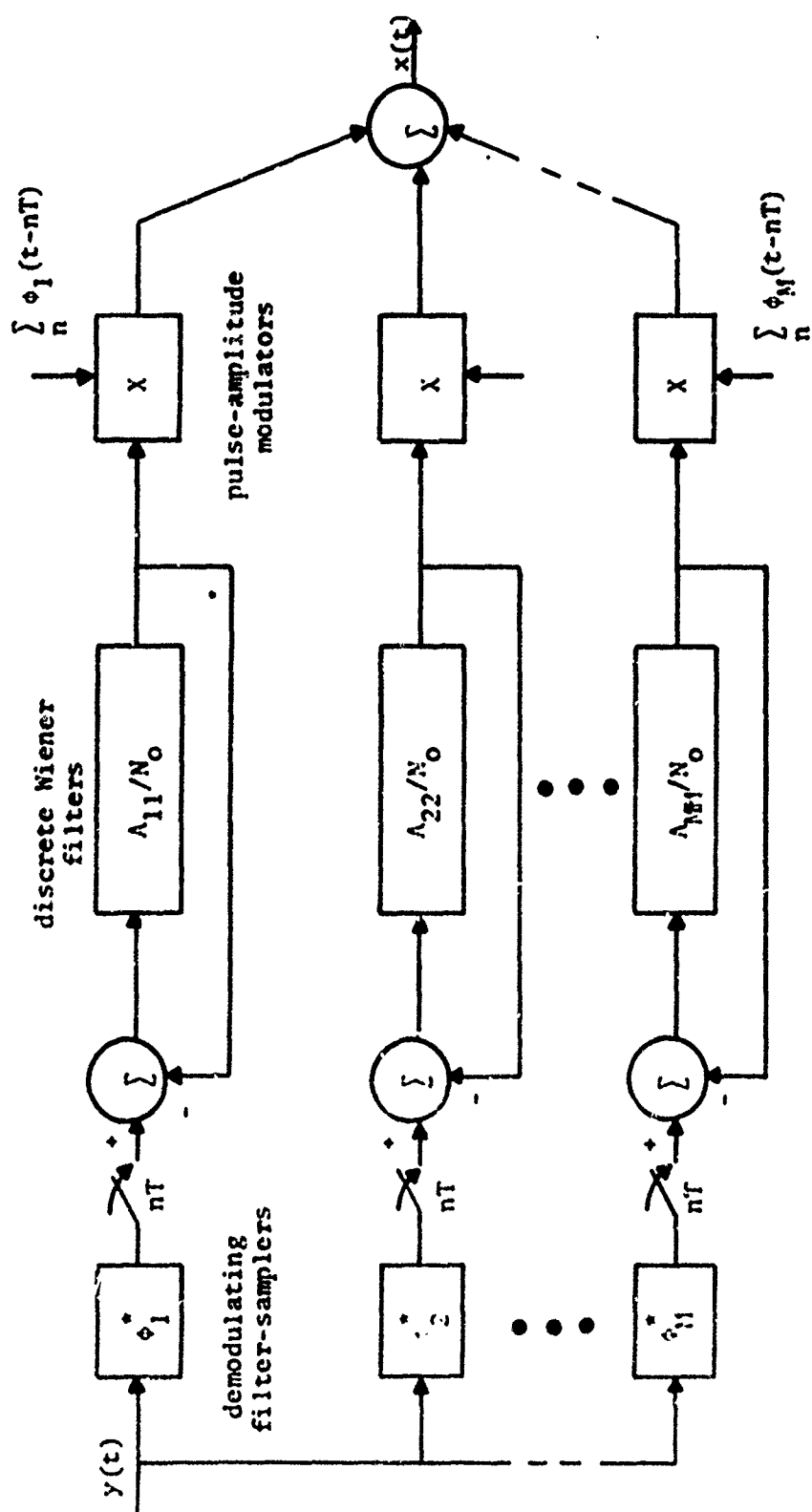
b)

Figure(4-13) a) Realization of the operator K_{yy} (T_1 , T_1^{-1} are periodic time-scale transformations and have impulse-response functions $\delta(g(t)-\tau)$, $\delta(t-g(\tau))$.) b) Realization of the operator K_{xy} . (For example 4.)



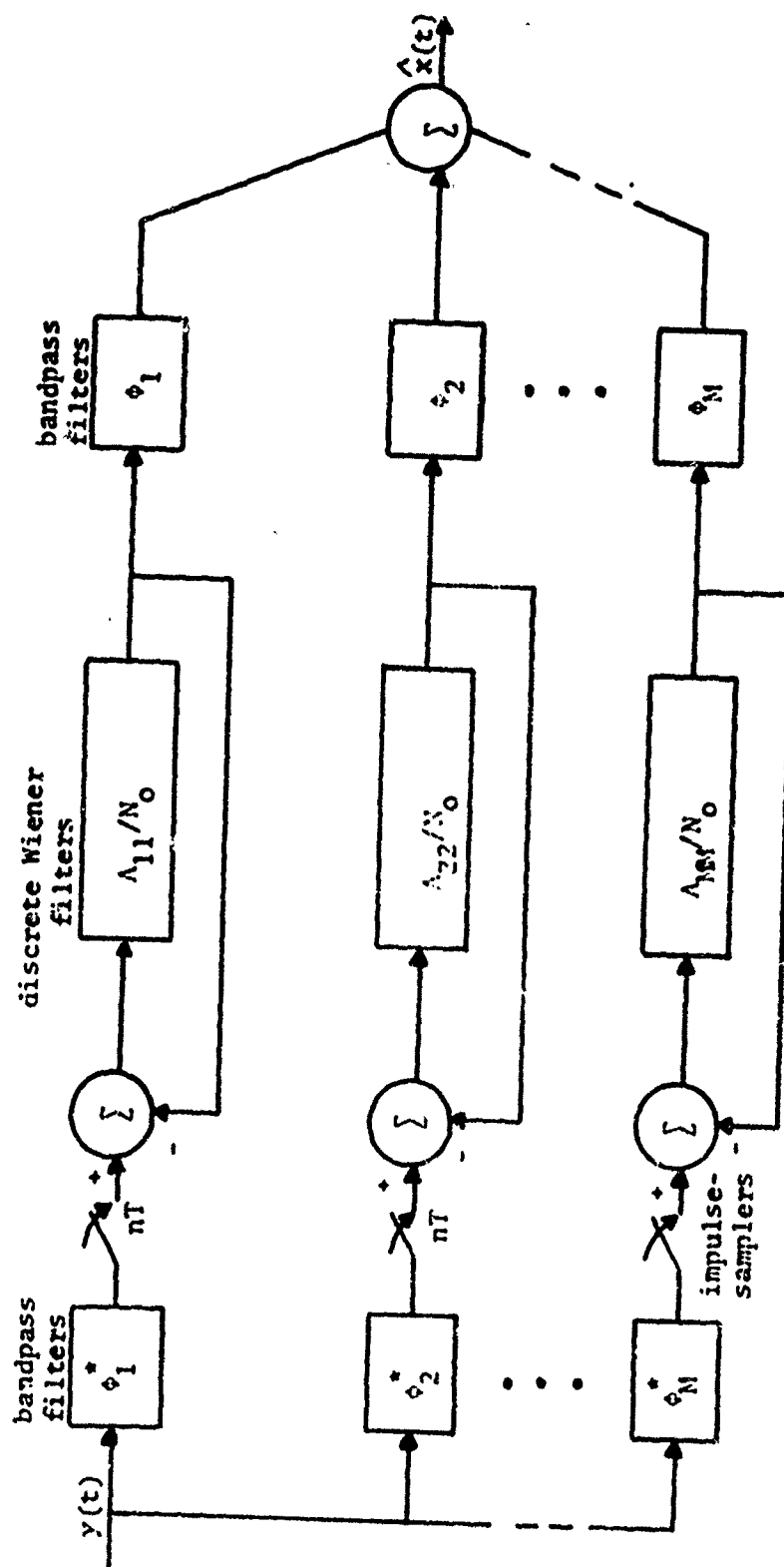
Figure(4-14) Structure of optimum time-varying filter for FSK signal; $\phi_p(t) = \sqrt{2B} \cos(2\pi p f_0 t) \frac{\sin(\pi B t)}{(\pi B t)}$,

$$\underline{H} = \underline{A}[\underline{A} + N \underline{I}]^{-1}, \quad A_{pq} = \frac{1}{M} \delta_{pq} - \frac{1}{M^2}.$$

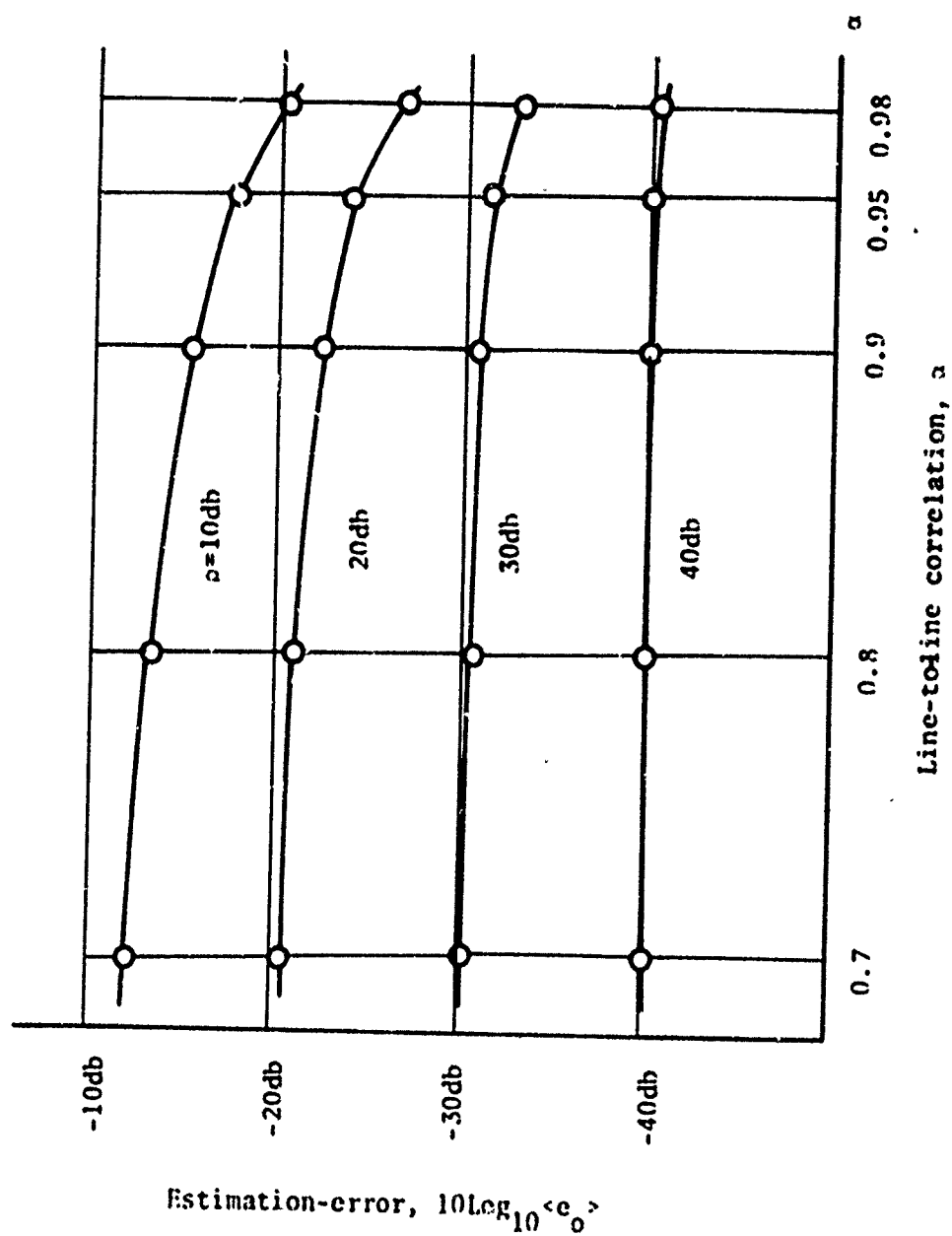


Figure(4-15) Structure of optimum time-varying filter for TDM-PM signal; $\phi_p(t) = \phi(t-pT/M)$,

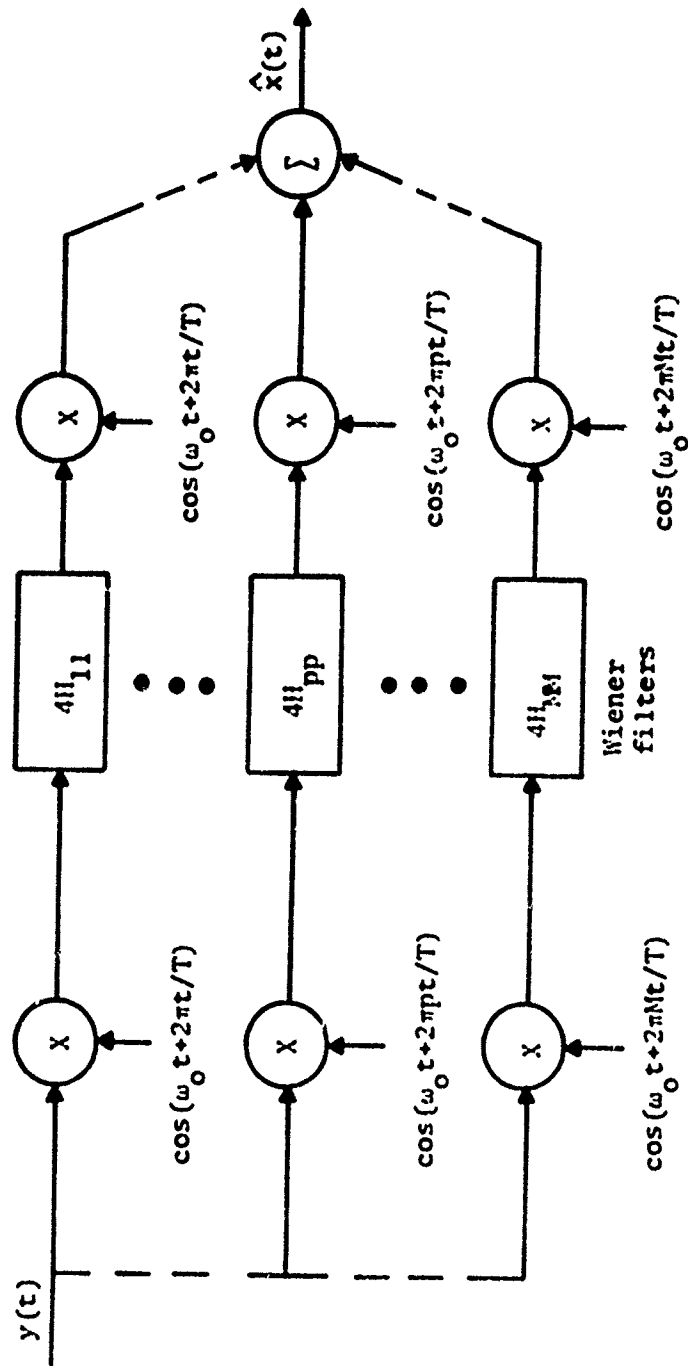
$$A_{pp}(f) = \frac{1}{M} \sum_q k_p(f-q/T).$$



Figure(4-16) Structure of optimum filter for FDM signal; $\Lambda_{pp}(f) = \frac{1}{2} \sum_q K_p(f-q/T)$, ϕ_p has center frequency $\omega_0/2\pi + p/T$, and bandwidth $1/T$. (See also Figure(4-18).)

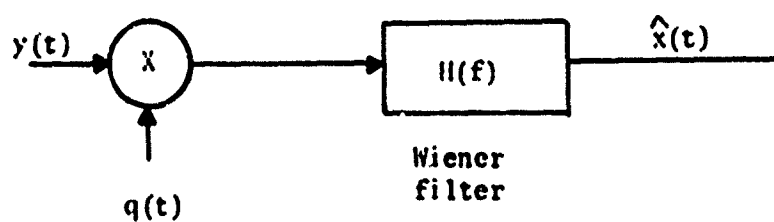


Figure(4-17) Time-averaged minimum estimation-error for video signal in additive white noise.

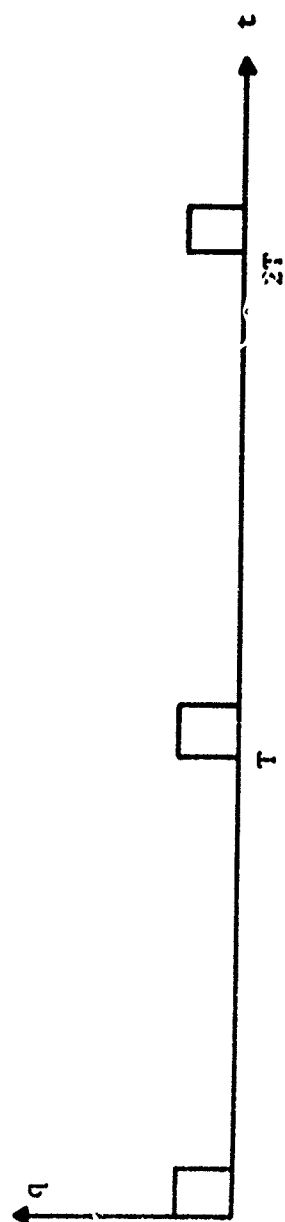


Figure(4-18) Structure of optimum time-varying filter for FDM signal;

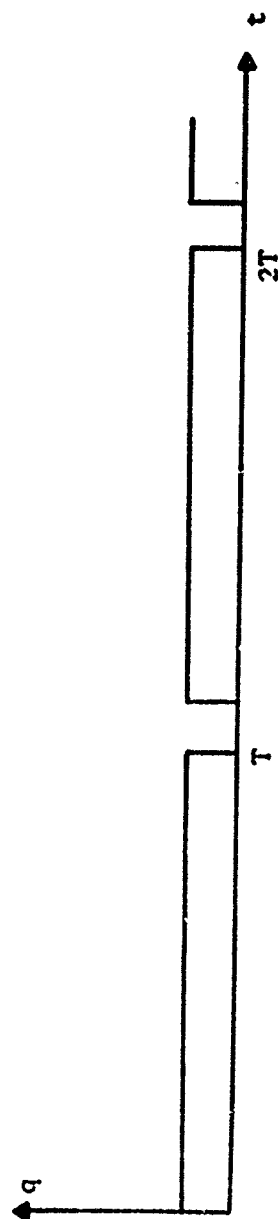
$$H_{pp}(\varepsilon) = K_p(\varepsilon) / (2K_p(\varepsilon) + 4N_0).$$



Figure(4-19) Optimum time-varying filter for periodically attenuated signal in additive white noise; $q(t) = 1 + \cos(2\pi t/T)$,
 $H(f) = K_{xx}(f) / (N_0 + (1 + \frac{\alpha^2}{2}) K_{xx}(f))$.

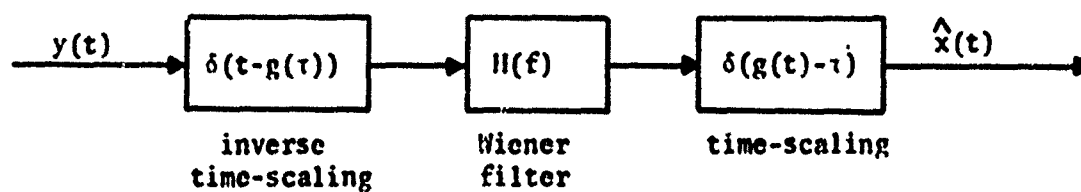


a)



b)

Figure(4-20) Gating functions.



Figure(4-21) Structure of optimum time-varying filter for periodically time-scaled processes; $H(f) = K_{zz}(f)/(P_0 K_{zz}(f) + N_0)$.

CHAPTER V

SUMMARY

In this final chapter we give a detailed outline and summary of the contents of Chapters II, III, IV which comprise the body of this dissertation. It is hoped that this combination index, outline, and summary will help to offset the length of Chapters II, III, IV. Following this summary, we suggest several topics for further research.

1. Summary of Chapters II, III, IV.

a) Chapter II: Transformation, Generation, and Modeling of Cyclostationary Processes. The second chapter of this dissertation provides a foundation upon which many of the developments in this dissertation are based, and upon which subsequent studies dealing with cyclostationary processes can be based. Specifically, Chapter II is devoted to the topics of generation of cyclostationary processes, and transformations on them. Characterizations of the various classes of transformations outlined below are developed and illustrated with numerous models for cyclostationary processes. These models, also listed below, are used throughout this dissertation for illustrating various theoretical results. Most of these models are of cyclostationary signals commonly used in communication and control systems.

Section 1 of Chapter II is introductory. In Section 2 we consider transformations which are linear. After deriving general input-output relations, we individually discuss, in some detail, the three subclasses of linear transformations referred to as time-invariant filters, periodically time-varying systems, and time-scale transformations. The impulse-response function and Zadeh's system function are introduced as means for characterizing linear systems, and are used throughout this section.

In Section 3, we briefly discuss multi-dimensional linear transformations as a means for characterizing scanning operations.

In the fourth section, we introduce a generalization of Wiener's Volterra series representation for characterizing input-output relations for periodic nonlinear systems.

Section 5 on random linear transformations parallels Section 2 on deterministic linear transformations. After deriving general input-output relations in terms of a generalization of Zadeh's system correlation function, we individually discuss, in some detail, the three subclasses of random linear transformations referred to as wide-sense stationary systems, cyclostationary systems, and random time-scale transformations.

Section 6 on random multidimensional linear transformations parallels Section 3. The major application of the theory in this section is to random scanning.

The final section on random nonlinear transformations is an extension of Section 4 on deterministic nonlinear transformations. The major issue in this section is jitter. This random disturbance is characterized in general terms, and shown to preserve cyclostationarity under fairly liberal conditions.

The specific contents of this chapter are outlined below:

Section 1: Introduction

Section 2: Linear Transformations.

a) Time-invariant filters.

Theorem (2-1): Preservation of cyclostationarity by time-invariant filters.

Theorem (2-2): Sampling theorem for nonstationary processes.

Theorem (2-3): Reduction of cyclostationary processes to stationary processes via bandlimiting filters.

b) Periodically time-varying systems.

Theorem (2-4): Generation of cyclostationary processes from stationary processes via periodic systems.

Model (1): Amplitude-modulated signals.

Model (2): Pulse-amplitude-modulated signals.

Model (3): Time-division-multiplexed signals.

Model (4): Frequency-division-multiplexed signals.

c) Time-scale transformations.

Theorem (2-5): Generation of cyclostationary processes from stationary processes via periodic time-scale transformations.

Model (5): Time-division-multiplexed signals.

Model (6): Doppler-shifted processes.

Model (7): Random facsimile signal.

Model (8): Asynchronous pulse-amplitude-modulated signals.

Section 3: Multi-dimensional linear transformations (scanning)

Theorem (2-6): Generation of one-dimensional cyclostationary processes from multi-dimensional stationary processes via multi-dimensional periodic systems.

Model (9): The video signal (line scanning).

Model (10): The passive radar signal (circular scanning).

Section 4: Nonlinear Transformations.

Theorem (2-7): Generation of cyclostationary processes from stationary processes via periodic nonlinear Volterra systems.

Model (11): Parametrically amplified signals.

a) Zero-memory nonlinearities.

Theorem (2-8): Generation of cyclostationary processes from stationary processes via periodic zero-memory nonlinearities.

b) Nonlinear modulation of synchronous pulse-trains.

Model (12): Frequency-shift-keyed signals (FSK).

Model (13): Phase-shift-keyed signals (PSK).

Model (14): Pulse-width-modulated signals (PWM).

Model (15): Pulse-position-modulated signals (PPM).

Model (16): Digital pulse-frequency-modulated signals (DPFM).

Model (17): Synchronous M-ary signals.

Theorem (2-9): Cyclostationarity of synchronous pulse-trains nonlinearly modulated with stationary sequences.

Section 5: Random linear Transformations.

a) Wide-sense stationary systems.

Theorem (2-10): Preservation of wide-sense stationarity by random transformations.

Model (18): Random periodic signals.

Theorem (2-11): Conversion of deterministic periodic systems into wide-sense-stationary systems via phase-randomization.

Theorem (2-12): Preservation of cyclostationarity by random transformations.

b) Cyclostationary systems.

Theorems (2-13): Generation of cyclostationary processes from
(2-14) stationary processes via random transformations.

Model (19): Random periodic attenuation of stationary processes.

Model (20): Multiplication of a stationary process by a cyclostationary process.

c) Random time-scale transformations.

i) Phase-randomization.

Theorem (2-15): Reduction of cyclostationary processes to stationary processes via phase-randomization

Theorem (2-16): Equivalence between phase-randomizing and time-averaging of random processes.

Theorem (2-17): Preservation of system cyclostationarity by phase-randomization.

Theorem (2-18): Reduction of cyclostationary systems to wide-sense stationary systems via phase-randomization.

Theorem (2-19): Equivalence between phase-randomizing and time-averaging random systems.

ii) Generalized angle modulation.

Theorem (2-20): Preservation of cyclostationarity by a stationary time-scale transformation.

Theorem (2-21): Generation of cyclostationary processes from periodic signals via stationary time-scale transformations.

Model (21): Phase-modulated signals.

Model (22): Frequency-modulated signals.

Theorem (2-22): Generation of cyclostationary processes from PAM time-scale transformations.

Model (23): Digital phase-modulated signals.

Model (24): Digital frequency-modulated signals.

Theorem (2-23): Generation of cyclostationary processes from stationary processes via cyclostationary time-scale transformations.

Section 6: Random Multi-dimensional Linear Transformations

Theorem (2-24): Generation of one-dimensional cyclostationary processes from multi-dimensional stationary processes via random multi-dimensional cyclostationary systems.

Section 7: Nonlinear Random Transformations.

a) Zero-memory random nonlinearities.

Theorem (2-25): Generation of cyclostationary processes from stationary processes via zero-memory random cyclostationary nonlinearities.

b) Jitter.

Theorem (2-26): Preservation of cyclostationarity of random processes by random jitter.

Theorem (2-27): Preservation of cyclostationarity of random systems by random jitter.

In concluding this summary of Chapter II, we mention two interesting facts which arise throughout the chapter: 1. There need be no exact periodicity in the realizations of a cyclostationary process: The occurrence of pulses in jittered PAM, for example, is not periodic, nor are the pulse-occurrences in a Poisson pulse-process with periodic rate parameter. The periodicity need only be present in the first and second order ensemble averages. 2. The mean and variance for a cyclostationary process can be constant so that the periodicity in these quantities is degenerate, and still such processes can exhibit strong periodic fluctuations in correlation, as exemplified by some time-division-multiplexed signals, pulse-amplitude-modulated signals, and others.

b) Chapter III: Series Representations for Cyclostationary Processes (and Their Autocorrelation Functions). The third chapter of this dissertation is an in-depth treatment of series representations for cyclostationary (and stationary) processes and their autocorrelation functions, and other periodic kernels; and includes discussions of their application to the solution of linear integral equations with periodic kernels, and to the realization of periodically time-varying linear systems, and to the definition of a generalized Fourier transform for cyclostationary processes.

Three types of series representations are treated:

1. Translation series representations of the form

$$x(t) = \sum_{p=1}^M \sum_{n=-\infty}^{\infty} a_{np} \phi_p(t-nT) \quad \forall t \in (-\infty, \infty)$$

$$k_{xx}(t, s) = \sum_{p, q=1}^M \sum_{n, m=-\infty}^{\infty} \Lambda_{n-m}^{pq} \phi_p(t-nT) \phi_q^*(s-mT) \quad \forall t, s \in (-\infty, \infty),$$

where $\{\phi_p\}$ are deterministic basis functions, and $\{a_{np}\}; p = 1, 2, \dots, M\}$ are jointly stationary random sequences with crosscorrelations

$$E\{a_{np} a_{mq}^*\} = \Lambda_{n-m}^{pq}.$$

(T is the period of cyclostationarity).

If the basis functions are doubly orthonormal in the sense that

$$\int_{-\infty}^{\infty} \phi_p(t-nT) \phi_q^*(t-mT) dt = \delta_{pq} \delta_{nm},$$

then the representors are given by the formulas:

$$a_{np} = \int_{-\infty}^{\infty} x(t+nT) \phi_p^*(t) dt$$

$$\Lambda_{n-m}^{pq} = \iint_{-\infty}^{\infty} k_{xx}(t+nT, \tau+mT) \phi_p^*(t) \phi_q(\tau) dt d\tau.$$

2. Harmonic series representations of the form:

$$x(t) = \sum_{p=-\infty}^{\infty} a_p(t) e^{j2\pi pt/T} \quad \forall t \in (-\infty, \infty)$$

$$k_{xx}(t, s) = \sum_{p, q=-\infty}^{\infty} k_{pq}(t, s) e^{j2\pi(p t - q s)/T} \quad \forall t, s \in (-\infty, \infty),$$

where the $\{a_p\}$ are jointly stationary processes bandlimited to $[-1/2T, 1/2T]$, and with crosscorrelations

$$E\{a_p(t)a_q^*(s)\} = k_{pq}(t-s).$$

The representors are given by the formulas:

$$a_p(t) = \int_{-\infty}^{\infty} w(t-\tau)x(\tau)e^{-j2\pi p\tau/T}d\tau$$

$$k_{pq}(t-s) = \iint_{-\infty}^{\infty} w(t-\tau)w(s-\gamma)k_{xx}(\tau,\gamma)e^{-j\pi(p\tau-q\gamma)/T}d\tau d\gamma$$

where

$$w(t) \triangleq \frac{\sin(\pi t/T)}{\pi t}.$$

3. Fourier series representations (for autocorrelation functions)

of the form

$$k_{xx}(t,s) = \sum_{n=-\infty}^{\infty} c_n(t-s)e^{j\pi n(t+s)/T} \quad \forall t,s \in (-\infty, \infty),$$

where the coefficient functions are defined to be

$$c_n(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} k_{xx}(t+\tau/2, t-\tau/2)e^{-j2\pi n\tau/T}d\tau.$$

These three types of representations are developed in detail, and the relationships existing amongst them are brought to light. The most general representation is the translation series representation, since many different translation series representations for a given cyclostationary process can be obtained by making different choices for the basis functions. In fact, a generalization of the Karhunen-Loeve representation is shown to be one special case of a translation series representation, and the Sampling Theorem provides another. It should be emphasized that these translation-series representations are countable discrete representations, and are valid on the entire real line $(-\infty, \infty)$.

The specific contents of this chapter are outlined below:

Section 1: Introduction

Section 2: Translation Series Representations.

a) Definition

Theorem (3-1): Existence and identification of translation series representations.

b) Representation in terms of PAM processes.

c) Implementation of the process-resolution operation.

d) Examples.

i) Harmonic translation series representation.

ii) Walsh translation series representation.

iii) Karhunen-Loeve translation series representation.

e) Finite-order translation series representations.

i) Bandlimited cyclostationary processes.

ii) Degenerate cyclostationary processes.

f) Correlation matrix decomposition and solution of integral equations.

i) Integral Equations.

Theorem (3-2): Explicit solution of general linear integral equation with periodic kernels.

ii) Correlation matrix decomposition.

(1) Infinite-dimensional diagonal matrices.

Theorem (3-3): Existence and identification of infinite-dimensional diagonal correlation matrices for translation series representations.

- (1a) Random video signal.
- (1b) Time-division-multiplexed signal.
- (2) Infinite-dimensional diagonal-plus-finite-rank matrices.

Theorem (3-4): Matrix inversion formula.

- (2a) Cyclostationary processes which are the outputs of periodically time-varying finite-dimensional linear dynamical systems driven by white noise.

Theorem (3-5): Existence and identification of diagonal-plus-finite-rank correlation matrix for translation series representation.

- (2b) Time-division-multiplexed process with component signals which have been generated by finite-dimensional time-invariant systems.

(3) Finite-dimensional diagonal and diagonal-plus-rank-1 matrices.

(3a) Time-division-multiplexed PAM signals.

(3b) Synchronous M-ary signals.

g) Implementation of periodically time-varying linear systems.

h) Generalized Fourier transform for cyclostationary processes.

Theorem (3-6): Equivalence between generalized Fourier transforms under statistical expectation (extension of Helly's Theorem).

Section 3: Harmonic Series Representation.

a) Definition.

Theorem (3-7): Existence and identification of harmonic series representation.

b) Resolution and reconstruction.

c) Properties of the harmonic series representations.

- (1) Stationarity.
- (2) Phase-randomization.
- (3) Bandlimitedness.
- (4) Time-invariant filtering.

d) Examples.

- (1) Amplitude-modulated signals.
- (2) Frequency-division-multiplexed signals.
- (3) Random periodic signals.

e) Solution of integral equations and realization of periodically time-varying systems.

Section 4: Fourier Series Representation for Autocorrelation Functions.

a) Definition.

Theorem (3-8): Identification of Fourier series coefficient functions and crosscorrelation functions for jointly stationary phase-randomized processes.

b) Properties of the Fourier series representation.

- i) Phase-randomization and stationarity.
- ii) Filtering and bandlimiting.
- iii) Constraints on coefficient functions.

c) Solution of integral equations and realization of periodically time-varying systems.

We conclude this summary of Chapter III by pointing out the fact that the essence of the representations presented here is that they effect a decomposition of a single cyclostationary process into a multiplicity

of jointly stationary components. Such representations enable one to convert many single-variable problems (involving cyclostationary processes) with periodically varying parameters to multi-variable problems with constant parameters. This is the case for optimum linear filtering problems as illustrated in Chapter IV.

Similarly, the autocorrelation-function representations presented here, when applied to the representation of impulse-response functions, effect a decomposition of a single periodically time-varying linear system into a multiplicity of time-invariant filters and periodic modulators.

c) Chapter IV: Least-Mean-Squared-Error Linear Estimation of Cyclostationary Processes. In this fourth chapter we consider a particular application of the theory developed in the preceding two chapters. These earlier chapters serve mainly as a theoretical base upon which various problems dealing with random periodic phenomena can be formulated and solved. Chapter II provides a base for obtaining mathematical models for physical systems and signals, so that physical problems can be formulated mathematically, and Chapter III provides a base for simplifying the mathematical models so that analyses can be simplified, and solutions obtained.

Although the application in this fourth chapter is but one of many possibilities, it has special significance in that it has been the primary motivation for developing the theory. Furthermore, the problem considered in this application is of significant practical (as well as theoretical) value in that its solution provides means for improving the

quality of present communication systems (and other signal processing systems), and a means for evaluating certain performance measures for newly proposed (as well as existing) communication systems, and it leads to useful insight into the nature of optimum signal processing.

Specifically, the problem considered is that of least-mean-squared-error linear estimation of imperfectly observed cyclostationary processes--or, in communications jargon, optimum filtering of noisy distorted cyclostationary signals. The particular aspects of the optimum filtering problem that are investigated in this chapter are the structure and the performance of optimum linear filters for cyclostationary signals. In fact, we prove (and demonstrate with numerous examples) the proposition that optimum filters for cyclostationary signals are periodically time-varying systems which can yield significantly improved performance over optimum filters which are constrained to be time-invariant and which thereby ignore the cyclic character of the cyclostationary signals. We also show that optimum linear receivers consist of demultiplexor-demodulators followed by optimum time-invariant filters--the order of these two operations being the reverse of that sometimes encountered in communication systems. These results are of considerable importance in view of the great preponderance of cyclostationary signals in communication systems (as emphasized throughout Chapter II). Specific examples which are worked out in detail are included in the outline below:

Section 1: Introduction

Section 2: The Orthogonality Condition.

Theorem (4-1): Lemma for the application of the Projection Theorem to the optimum filtering problem.

Theorem (4-2): Statement of the orthogonality condition as it applies to the optimum filtering problem.

(1) "Infinite-dimensional" filters.

(2) "Finite-dimensional" filters.

Section 3: Optimum Time-Invariant Filters for Stationary Processes.

Theorem (4-3): Bounds on mean-squared estimation-error.

Section 4: Optimum Time-Invariant Filters for Cyclostationary Processes.

Theorem (4-4): Equivalence between the optimum filters which minimize the time-averaged error for cyclostationary processes, and those which minimize the constant error for the phase-randomized versions of the cyclostationary processes.

Section 5: Optimum Time-Varying Filters for Cyclostationary Processes.

a) Introduction.

b) Solutions based on translation series representations.

Theorem (4-5): General solution for the impulse-response function for the optimum time-varying filter for cyclostationary processes, and solution for the minimum estimation-error.

Theorem (4-6): Extension of Theorem (4-5).

Theorem (4-7): Special case of Theorems (4-5), (4-6).

Theorem (4-8): Bounds on mean-squared estimation-error.

c) Solutions based on the harmonic series representation.

i) Estimation of cyclostationary processes on an infinite interval.

Theorem (4-9): General solution for the impulse-response function for the optimum time-varying filter for cyclostationary processes, and solution for the minimum estimation-error.

ii) Estimation of periodic processes on a finite interval.

Theorem (4-10): Solution for the impulse-response function for the optimum time-varying smoother for periodic processes in additive white noise, and solution for the minimum estimation-error.

Theorem (4-11): Counterpart of Theorem (4-10) for the optimum time-invariant smoother.

d) Solutions based on the Fourier series representation for correlation functions.

Theorem (4-12): Conversion of integral equation for the impulse-response function for the optimum time-varying filter to an infinite set of linear algebraic equations.

Theorem (4-13): Specific case of Theorem (4-12) which yields an explicit solution.

e) Solutions based on inverse-operator and feedback-system methods.

i) Introduction.

ii) Examples.

(1) WSS signal subjected to time-invariant dispersion and periodic attenuation, and additive white noise.

- (2) CS AM signal in additive WSS white noise.
- (3) WSS signal in additive CS white noise.
- (4) CS signal (obtained from periodically time-scaled WSS process) in additive WSS white noise.

Section 6: Examples of Optimum Filters and Improvements in Performance.

a) Examples employing translation series representations.

- i) Frequency-shift-keyed signals.
- ii) Time-division-multiplexed PAM signals.
- iii) Frequency-division-multiplexed signals.
- iv) Pulse-amplitude-modulated signals.
- v) PAM signals through random channels.
- vi) The video signal.
- vii) Time-division-multiplexed signals.

b) Examples employing the harmonic series representation.

- i) Frequency-division-multiplexed signals.
- ii) Amplitude-modulated signals.
- iii) The random sinusoid.

c) Examples employing the Fourier series representation.

- i) Amplitude-modulated signals.
- ii) Bandpass PAM signals.
- iii) Improvement for general PAM.

d) Examples employing the inverse-operator-feedback-system method.

- i) Periodically attenuated WSS signal in WSS noise.
- ii) AM-type signal in WSS noise.
- iii) WSS signal in white CS noise.
- iv) Doppler-shifted Poisson pulse-train.
- v) Time-division-multiplexed signals.

f) Summary.

Theorem (4-14): Demultiplexing and demodulation precedes time-invariant filtering in optimum linear receiver.

We conclude this summary of Chapter IV by mentioning the intuitively pleasing result that optimum time-varying filters for cyclostationary processes can always be decomposed into a sequence of three signal processing operations: 1. The received cyclostationary process is decomposed into one or more (jointly) stationary processes. 2. The stationary component-processes are passed through time-invariant filters (often Wiener filters). 3. The filtered components are recombined to form the optimally filtered cyclostationary process. In most of the examples presented, this sequence of operations takes one of the following specific forms: demodulator-filter-modulator, demultiplexor-filter-multiplexor, inverse time-scale transformation--filter--time-scale transformation. For those particular cyclostationary processes which decompose into a single stationary process, the three signal processing operations are more accurately described as: 1. Removal of periodic fluctuations. 2. Time-invariant filtering. 3. Re-insertion of periodic fluctuations.

Another result worth emphasizing is that cyclostationary processes whose mean and variance are constant can still possess a high degree of cyclostationarity in the sense that optimum time-varying filtering can yield substantially lower mean-squared error than optimum time-invariant filtering.

Finally, we mention that optimum time-varying filters for cyclostationary processes appear to yield the greatest reduction in mean-squared error (below that resulting from optimum time-invariant filters) when the signal-to-noise ratio is high, but can yield substantial reductions even at signal-to-noise ratios as low as unity.

2. Suggested Topics for Further Research

There are various topics dealing with cyclostationary processes which deserve further research. We outline below several such areas of investigation.

a) Degree of cyclostationarity. The only effective means we have for determining the filtering advantage to be gained from a periodically time-varying filter operated in synchronism with the received cyclostationary signal is to carry out the complete solutions for the optimum time-varying and time-invariant filters, and to evaluate the performance functionals. It would therefore be desirable to devise a functional which will indicate the "degree of cyclostationarity" of a process with regard to filtering-performance improvement that inherently can be attained. Obviously, a "highly cyclostationary" process will exhibit the largest improvement. One might expect that a constant-variance cyclostationary process must have a limited degree of cyclostationarity. The video signal, for example has constant variance and for typical parameter values the performance improvement factor is essentially negligible. However, the PAM signal with full duty-cycle rectangular pulses has constant variance, yet has an unlimited improvement factor for low-noise conditions. One possible candidate for measuring the degree of cyclostationarity is the relative size of the off-diagonal terms in the HSR correlation matrix. For example, if the off-diagonal terms vanish then the process is stationary, and by definition there is no improvement. A relatively simple functional for indicating degree of cyclostationarity or obtaining bounds on improvement should find extensive application in

further studies since certain aspects of performance can be determined without having to carry out a complete solution to the estimation problem.

b) Implementation of periodically time-varying filters. Although most of the examples of optimum filters presented in Chapter IV were finite structures, the general solutions provided in Theorems (4-5,6,7,9) are based on infinite series representations, and result in infinite structures (infinite number of parallel paths.) As discussed in Chapter III, truncated versions of the representations may provide satisfactory approximations and finite structures, and the implementation of such structures can be enhanced by the use of convenient basis functions which result in the replacement of modulators by switches, and by the use of digital filters to implement sampled-data filters. However, these filters probably do not have the minimum complexity for a given degree of approximation to the optimum filter. Less complex filter structures might likely result from an a-priori choice of basis filters, and solution for the corresponding optimum "finite-dimensional filter" as discussed in Section 2 of Chapter IV. In fact, the basis filters can be chosen for adjustability so that changes in signal and noise parameters can be conveniently accommodated. An investigation into the trade-off's available between convenience-of-implementation and performance could produce some useful results.

c) Synchronizing and tracking. A random signal can be cyclostationary even though there may be no strict periodicity in the occurrence of the random events comprising the signal. For example, in direct contrast to

the periodic occurrence of pulses in a synchronous PAM signal, the pulse-occurrences in a Poisson pulse-train process with periodic rate parameter are random--only the average rate-of-occurrences is periodic. In fact some signal processes exhibit cyclic fluctuations but are only "quasi-cyclostationary" since the period of the cyclic fluctuations varies slowly with time. Obtaining knowledge of the phase of such non-synchronous cyclostationary (and quasi-cyclostationary) signals is not as straightforward as it is for highly structured synchronous cyclostationary signals such as PAM. Thus it would be desirable to develop methods for tracking cyclostationary signals in order to extract useful timing information for optimum filtering, and other signal processing operations. The phase lock loop and some of its variations such as the Costas filter loop and the squaring loop have found much application for symbol synchronization for synchronous pulse-trains, and for carrier synchronization [40]. One approach to the timing recovery problem for more general cyclostationary signals then would be to extend these and related ideas which have been successful for the more structured signals [1,40]. Future investigations might seek to evaluate the effects of trade-offs in complexity between the time-varying filter and the tracking-loop structures on overall performance.

d) Modelling. A continued study of appropriate models for random processes resulting from various types of sampling, scanning, and multiplexing operations is essential in order to take full advantage of the results in this thesis. In particular, one might concentrate on operations

involving random parameters, such as scanning speed or clocking rate, which produce non-synchronous signals. The resulting processes may be either cyclostationary or "quasi-cyclostationary", as mentioned previously, depending on the models chosen. See, for example, Section 7 of Chapter II on jitter.

e) Signal Design. Related to the modelling problem above are questions concerning the design of the sampling, scanning, and multiplexing operations with regard to introducing a sufficient degree of cyclostationarity for satisfactory receiver operation, with respect to both timing extraction and filter performance. An investigation of such ideas could provide the basis for the joint optimum design of transmitter and receiver. Previous studies have treated an aspect of this known as the power allocation problem [40] wherein the trade-off between power in synchronizing signals and in information-bearing signals is examined. A new approach to this problem might be based on degree of cyclostationarity.

BIBLIOGRAPHY

1. Bennett, W.R., "Statistics of regenerative digital transmission," Bell Sys. Tech. Jour., vol. 37, pp. 1501-42, November 1958.
2. Franks, L.E., Signal Theory, Prentice-Hall, 1969.
3. Papoulis, A., Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1965.
4. Gladyshev, E.G., "Periodically and almost periodically correlated random processes with continuous time parameter," Theory of Prob. and its Appl., pp. 173-177, 1963.
5. Hurd, H., "An investigation of periodically correlated stochastic processes," Ph.D. Dissertation, Duke University, 1969.
6. Ogura, H., "Spectral representation of a periodic nonstationary random process," IEEE Trans. Inform. Theory, vol. IT-17, pp. 143-149, March 1971.
7. Stratonovich, R.L., Topics in the Theory of Random Noise, Gordon and Breach, 1963.
8. Deutsch, R., "Detection of modulated noise-like signals," IEEE Trans. on Info. Theory, 1954, pp. 106-122.
9. Bendat, J.S. and A.G. Piersol, Measurement and Analysis of Random Data, New York: Wiley, 1966.
10. Gardner, W.A., "The structure of linear L.M.S. estimators of cyclostationary synchronous m-ary signals," Proc. of the Sixth Annual Princeton, Conf. on Info. Sci. and Sys., 1972.

11. Franks, L.E., "A model for the random video process,"
Bell Sys. Tech. Jour., vol. 45, April 1966, pp. 609-630.
12. Macfarlane, G.G., "On the energy spectrum of an almost periodic succession of pulses," Proc. I.R.E., vol. 37, October 1949, pp. 1139-1143.
13. Rice, S.O., "Mathematical analysis of random noise," Bell Sys. Tech. Jour., vol. 23, July 1944, pp. 282-332.
14. Gudzenko, L.I., "On periodic nonstationary processes," Radio Engineering and Electronic Physics, vol. 4, No. 6, 1959, pp. 220-224 (trans.).
15. Gladyshev, E.G., "Periodically correlated random sequences," Soviet Mathematics, vol. 2, 1961, pp. 385-388 (trans.).
16. Oliver, B.M., J.R. Pierce and C.E. Shannon, "The philosophy of PCM," Proc. I.R.E., vol. 36, November 1948, pp. 1324-1331.
17. Monin, A.S., "Stationary and periodic time series in the general circulation of the atmosphere," Proc. of the Symp. on Time Series Anal., 1962, Ed. by M. Rosenblatt, New York: Wiley, pp. 144-151, 1963.
18. Willis, D.M., "The statistics of a particular non-homogeneous Poisson process," Biometrika, vol. 51, 1964 (Great Britain) pp. 399-404.
19. Brelsford, W.M., "Probability predictions and time series with periodic structure," Ph.D. Dissertation, Johns Hopkins Univ. 1967.
20. Jones, R.H. and W.M. Brelsford, "Time series with periodic structure," Biometrika, vol. 54, 1967, pp. 403-407.
21. Markelov, V.A., "Axis crossings and relative time of existence of a periodically nonstationary random process," Soviet Radiophysics, 1966, pp. 440-443.

22. Parzen, E., Stochastic Processes, San Francisco: Holden-Day, 1962, p. 157.
23. Anderson, R.R., and J. Salz, "Spectra of digital FM," Bell Sys. Tech. Jour., vol. 44, August 1965, pp. 1165-1189.
24. Lundquist, L., "Digital PM spectra by transform techniques," Bell Sys. Tech. Jour., vol. 48, February 1969, pp. 397-441.
25. Jordan, K.L., "Discrete representations of random signals," Ph.D. Dissertation, Mass. Inst. of Tech. (also MIT RLE Tech. Report 378). July 1961.
26. Bartle, R.G., The Elements of Real Analysis, Wiley, 1964.
27. Kingman, J.F.C., and S.J. Taylor, Introduction to Measure and Probability, Cambridge, 1966.
28. Wiener, N., The Fourier Integral and Certain of its Applications, (Cambridge, 1933) Dover, 1958.
29. Epstein, B., Linear Functional Analysis, Saunders, 1970.
30. Zemanian, A.H., Distribution Theory and Transform Analysis, McGraw-Hill, 1965.
31. Riesz, F., and B. Sz-Nagy, Functional Analysis, (1952) Frederick Ungar, 1965.
32. Titchmarsh, E.C., The Theory of Functions, Oxford (1932), 1964.
33. Titchmarsh, E.C., The Theory of Fourier Integrals, Oxford (1937), 1967.
34. Pugachev, V.S., Theory of Random Functions, (1948), Addison-Wesley 1965.
35. Zadeh, L.A., "Frequency analysis of variable networks," Proc. I.R.E., vol. 38, 1950, pp. 291-299.

36. D'Angelo, H., Linear Time-Varying Systems: Analysis and Synthesis, Allyn and Bacon, 1970.
37. Zadeh, L.A., "Correlation functions and power spectra in variable networks," Proc. I.R.E., vol. 38, November 1950, pp. 1342-1345.
38. Maurer, R.E., "The optimal equalization of random channels," Ph.D. Dissertation (also Comm. Th. Group Report No. 9), Northeastern University, 1968.
39. Hartley, G.C., P. Mornet, F. Ralph, D.J. Tarran, Techniques of pulse-code-modulation in communication networks, Cambridge 1967.
40. Stifler, J.J., Theory of Synchronous Communications, Prentice-Hall, 1971.
41. Middleton, D., "A statistical theory of reverberation and similar first-order scattered fields"--Parts II and IV, IEEE Trans. on Info. Th., vols. IT-13,18, July 1967, pp. 393-414 and Jan. 1972, pp. 68-90.
42. Middleton, D., "A statistical theory of reverberational and similar first-order scattered fields"--Part III, IEEE Trans. on Info. Th., vol. IT-18, January 1972, pp. 55-67.
43. Franks, L.E., and W.A. Gardner, "Estimation for cyclostationary random processes," Proc. of Ninth Annual Allerton Conf. on Ckt. and Sys. Th., 1971, pp. 212-231.
44. Wiener, N., Nonlinear Problems in Random Theory, Wiley, 1958.
45. Deutsch, R., "On a method of Wiener for noise through nonlinear devices," IRE Convention Record, pt. 4, 1955, pp. 186-192.
46. George, D.A., "Continuous nonlinear systems," M.I.T.-R.I.E. report no. 355, July, 1958.

47. Bedrosian, E., and S.O. Rice, "The output of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs," Proc. of IEEE, vol. 59, December 1971, pp. 1688-1707.
48. Stern, T.E., Theory of Nonlinear Networks and Systems, Addison-Wesley, 1965.
49. Sage, A.P. and J.L. Melsa, Estimation Theory with Applications to Communications and Control, McGraw-Hill, 1971.
50. Meditch, J.S., Stochastic Optimal Linear Estimation and Control, McGraw-Hill, 1969.
51. Van Trees, H.L., Detection, Estimation and Modulation Theory, Wiley, 1968.
52. Brockett, R.W., Finite Dimensional Linear Dynamical Systems, Wiley, 1970.
53. Kuo, B.C., Discrete Data Control Systems, Prentice-Hall, 1970.
54. Jury, E.I., Theory and Application of the Z-Transform Method, Wiley, 1964.
55. Papoulis, A., The Fourier Integral and its Applications, McGraw-Hill, 1962.
56. Wiener, N., The Extrapolation, Interpolation, and Smoothing of Stationary Time Series, Wiley, 1949.
57. Zadeh, L.A., and J.R. Ragazzini, "An extension of Wiener theory of prediction," Jour. Appl. Phys., vol. 21, 1950, pp. 645-655.
58. Booton, R.C., "An optimization theory for time-varying linear systems with nonstationary statistical inputs," Proc. I.R.E., vol. 40, 1952, pp. 977-981.

59. Luenberger, D.G. Optimization by Vector Space Methods, Wiley, 1969.
60. Kalman, R.E., "A new approach to linear filtering and prediction problems," Trans. ASME Jour. of Basic Eng., vol. 82D, March 1960, pp. 35-45.
61. Kalman, R.E., and R.S. Bucy, "New results in linear filtering and prediction theory," Trans. ASME Jour. on Basic Eng., March 1961, vol. 83D.
62. Gardner, W.A., "Fundamental results on the structure of linear estimators with applications to estimation and detection of synchronous m-ary signals," IEEE Trans. on Information Theory, to appear.
63. Maurer, R.E., and L.E. Franks, "Optimal linear processing of randomly distorted signals," IEEE Trans. on Circuit Theory, vol. CT-17, February 1970, pp. 61-67.